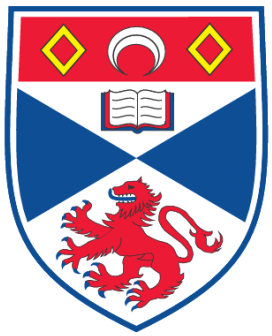
BALANCE, GRAVITY WAVES AND JETS IN TURBULENT SHALLOW WATER FLOWS

Jemma Shipton

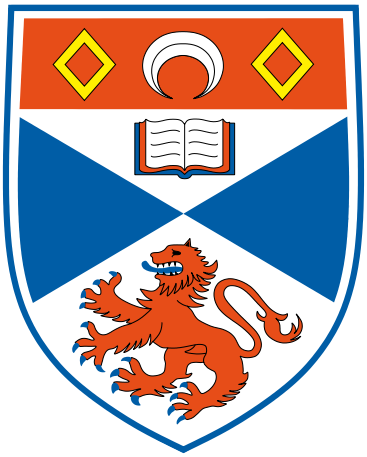
A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews



2009

Balance, Gravity Waves and Jets in Turbulent Shallow Water Flows.

Jemma Shipton



A thesis submitted for the degree of Doctor of Philosophy at the University of St Andrews

July 2008

Abstract

This thesis contains a thorough investigation of the properties of freely decaying turbulence in a rotating shallow water layer on a sphere. A large number of simulations, covering an extensive range of Froude and Rossby numbers, have been carried out using a novel numerical algorithm that exploits the underlying properties of the flow. In general these flows develop coherent structures; vortices interact, merge and migrate polewards or equatorwards depending or their sign, leaving behind regions of homogenized potential vorticity separated by sharp zonal jets. In the first half of the thesis we investigate new ways of looking at these structures. In the second half of the thesis we examine the properties of the potential vorticity (PV) induced, balanced component and the residual, unbalanced component of the flows.

Cyclone-anticyclone asymmetry has long been observed in atmospheric and oceanic data, laboratory experiments and numerical simulations. This asymmetry is usually seen to favour anticyclonic vorticity with the asymmetry becoming more pronounced at higher Froude numbers (e.g. Polvani et al. [1994a]). We find a similar result but note that the cyclones, although fewer, are significantly more intense and coherent. We present several ways of quantifying this across the parameter space.

Potential vorticity homogenization is an important geophysical mechanism responsible for sharpening jets through the expulsion of PV gradients to the edge

of flow structures or domains. Sharp gradients of PV are obvious in contour plots of this field as areas where the contours are bunched together. This suggests that we can estimate the number of zonal jets by performing a cluster analysis on the mean latitude of PV contours (this diagnostic is also examined by Dritschel and McIntyre [2007]). This provides an estimate rather than an exact count of the number of jets because the jets meander significantly. We investigate the accuracy of the estimates provided by different clustering techniques. We find that the properties of the jets defy such simple classification and instead demand a more local examination. We achieve this by examining the palinstrophy field. This field, calculated by taking the gradient of the PV, highlights the regions where PV contours come closer together, exactly what we would expect in regions of strong jets. Plots of the palinstrophy field reveal the complex structure of these features.

The potential vorticity field is even more central to the flow evolution than the strong link with jets suggests. From a knowledge of the spatial distribution of PV, it is possible to *diagnose* the balanced components of all other fields. These components will not contain inertia-gravity waves but will contain the dominant, large scale features of the flow. This inversion, or decomposition into balanced (vortical) and unbalanced (wave) components, is not unique and can be defined to varying orders of accuracy. We examine the results of four different definitions of this decomposition, two based on truncations of the full equations and two based on an iterative procedure applied to the full equations. We find the iterative procedure to be more accurate in that it attributes more of the flow to the PV controlled, balanced motion. However, the truncated equations perform surprisingly well and do not appear to suffer in accuracy at the equator, despite the fact that the scaling on which they are based has been thought to break down there.

We round off this study by considering the impact of the unbalanced motion on

the flow. This is accomplished by splitting the integration time of the model into intervals $\tau < t < \tau + \mathrm{d}\tau$ and comparing, at the end of each interval, the balanced components of the flow obtained by a) integrating the model from $t = 0$ and b) integrating the full equations, initialised at $t = \tau$ with the balanced components from a) at $t = \tau$. We find that any impact of the unbalanced component of the flow is less than the numerical noise of the model.

Acknowledgements

I would like to thank my supervisor, David Dritschel, for his guidence, support and wonderful cuisine. I have thoroughly enjoyed my time at St Andrews and I feel privileged to have been part of the vortex dynamics research group.

I would also like to thank my husband, Daniel Harding, and my family for their unwavering support throughout my studies.

This research was funded by NERC (Grant F14/G6/109).

I, Jemma Shipton, hereby certify that this thesis, which is approximately 20,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.
I was admitted as a research student in October 2003 and as a candidate for the degree of PhD in October 2003; the higher study for which this is a record was carried out in the University of St Andrews between 2003 and 2006.

Date 13/02/09 Signature of candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 26/02/09 Signature of Supervisor

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Access to Printed copy and electronic publication of thesis through the University of St Andrews.

Date 13/02/09

Signature of candidate

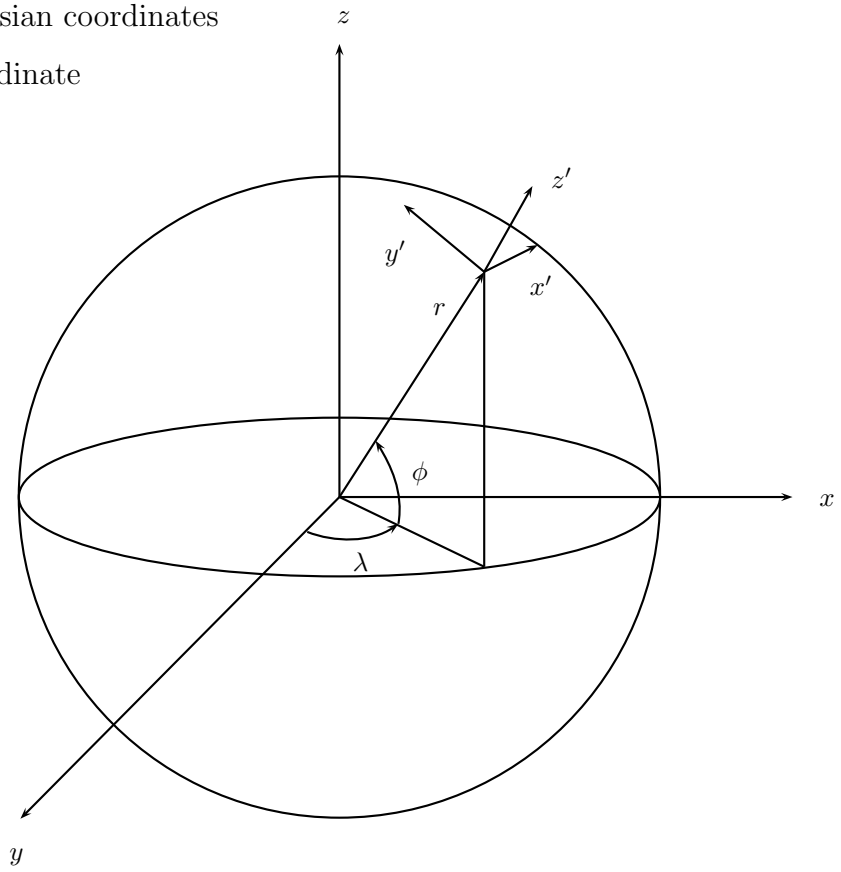
Signature of supervisor

Contents

Commonly used symbols

0.1 Spherical geometry

x, y, z	Cartesian coordinates
x', y', z'	local Cartesian coordinates
r	radial coordinate
ϕ	latitude
λ	longitude



0.2 Fluid dynamics

Π	potential vorticity
$\boldsymbol{\omega}$	relative vorticity
ρ	density
$\boldsymbol{u} = (u, v, w)$	fluid velocity
$\boldsymbol{\Omega}$	angular velocity of reference frame
p	pressure
f	Coriolis parameter
β	latitudinal variation of the Coriolis parameter
g	acceleration due to gravity
h	fluid depth
H	mean fluid depth
ζ	vertical component of the relative vorticity
$\delta = \nabla \cdot \boldsymbol{u}$	divergence of fluid velocity
$\gamma = \nabla \cdot \boldsymbol{a} \equiv \nabla \cdot \dfrac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}$	divergence of fluid acceleration

0.3 Dimensionless parameters

See table 1.1

0.4 Clustering parameters

See table 3.3

0.5 Abbreviations

PV	potential vorticity
CASL	contour-advective semi-Lagrangian
QG	quasi-geostrophic
OPV	optimal potential vorticity

Chapter 1

Introduction

> *We might say that the atmosphere is a musical instrument on which one can play many tunes. High notes are sound waves, low notes are long inertial waves, and nature is a musician more of the Beethoven than of the Chopin type. He much prefers the low notes and occasionally plays arpeggios in the treble and then only with a light hand.*
> From an unpublished letter from Jule Charney to Phillip Thompson, 12 February 1947

> *Extending Charney's analogy, we can think of inertia-gravity waves as meteorological noise that is not loud enough to drown out the symphony orchestra though it does detract from the performance. In the atmosphere, this noise is an inconvenience; in numerical models, it can be disastrous.*
> Daley, 1991

The quotations above provide an analogy that gives an intuitive insight into the central theme of this thesis: the concept of balance in turbulent geophysical flows. A balanced flow is Daley's symphony orchestra or Charney's low notes.

Superimposed on this balanced, large scale, slow motion are the inertia-gravity waves, the noise or, more poetically, the 'treble arpeggios'. As indicated by the musical analogy there is, to some extent, a frequency separation between these two types of motion. Observations show that the large scale motions of both the atmosphere and oceans are close to being balanced. Inertia-gravity waves are present but can, in many situations, be considered to have little impact on the flow. So why can we not neglect them entirely? Attempts have been made to do this by filtering the equations so that such waves are not permitted solutions. These simplified models have greatly aided understanding of important dynamical process and are at least qualitatively applicable to many atmospheric and oceanic systems. However, filtering the waves can have a detrimental effect on the remaining balanced, or vortical, component of the flow. This could either be because the waves themselves are important, as they undoubtedly are in some regions of the atmosphere and oceans, or it could be because the frequency separation between the balanced (vortical) and unbalanced (inertia-gravity wave) components of the flow is not sufficiently precise. If it is the second case then the model in question may be improved by returning to the original equations and retaining higher order terms. However, this presupposes that there *is* a balanced flow that can be achieved. This is not the case. Flows that are initially balanced will not necessarily remain balanced and spontaneous emission of inertia-gravity waves can and does occur. To return to the musical analogy, the distinction between high and low notes only tells part of the story. The low notes in fact contain higher harmonics which, although they have minimal effect on the pitch of the note, influence the tone and texture of the music.

Two questions arise out of this musical analogy: firstly, how accurately can the balanced flow be determined; and secondly, given the inherent inaccuracy of such a decomposition, to what extent is the concept of balance of use? The current status of research on the first question will be briefly outlined later in the

introduction and we shall return to the question in chapter 4. To motivate this we shall now turn to the second question with a discussion of the importance of balance and the many ways in which the concept can be exploited to obtain an insight into fundamental aspects of fluid motion.

1.1 Balance

In the previous section we introduced, via a musical analogy, the idea that two fundamentally different types of motion exist in geophysical fluid flows (see figure 1.1 for examples of these two types of motion). In these flows it is the low frequency 'slow' or 'vortical' motions that dominate while the higher frequency 'fast' or 'inertia-gravity wave' oscillations are generally observed to have a much smaller amplitude. This means that geophysical flows are close to being 'balanced flows'. The definition of a balanced flow is subtle. If a balanced flow is considered to be one where the forces acting on a fluid parcel are balanced then clearly there can be no motion. In reality the forces acting on a fluid parcel are only *close* to being in balance and this balance is revealed through an analysis of the order of magnitude of the terms in the equations of motion. Equating the zeroth order terms gives diagnostic balance relations but in order to permit motion the first order terms at least have to be retained (see section 2.3). The balanced model then comprises a prognostic equation for the time evolution of the 'master variable' and a set of balance relations taking the form of diagnostic equations that calculate the other variables. Typically the 'master variable' is based on vorticity which is why 'vortical motion' and 'balanced motion' are used interchangeably. The link between the master variable and the balanced flow is clear if the equations are linearised about a state of rest (see section 2.2). However, the equations do not have to be linearised for balance relations to be defined. Instead it is possible to expand the equations in terms of dimensionless

parameters and truncate at some particular order. Yet another approach is to filter the inertial gravity waves from a time series of data. This involves implicit assumptions about the frequency separation between the balanced and 'unbalanced' components of the flow. As indicated by the analogy above this separation is not clear. Consequently this approach results in a significant portion of the balanced flow being diagnosed as unbalanced. In the second part of this thesis we apply several different definitions of balance to the turbulent flows considered in the first part of the thesis in order to investigate the extent to which flows remain balanced as they evolve.

In balanced models of geophysical flows the usual 'master variable' is the potential vorticity (PV). The potential vorticity, denoted by Π, is a scalar field that is proportional to the dot product of the vorticity and the gradient of a materially conserved quantity:

$$\Pi = \frac{\boldsymbol{\omega} + 2\boldsymbol{\Omega}}{\rho} \cdot \boldsymbol{\nabla}\lambda, \tag{1.1}$$

where ρ is the fluid density, $\boldsymbol{\omega}$ is the relative vorticity, $\boldsymbol{\Omega}$ is the angular velocity of the reference frame and λ is the materially conserved quantity [Pedlosky, 1987]. A balanced flow is then a flow for which all the dynamical information is contained in the potential vorticity field. This is known as the 'invertibility principle' [Hoskins et al., 1985].

The concept of balance is important because it allows us to simplify the equations for the evolution of the flow. Taking the material derivative of equation 1.1 in the absence of friction and under certain constraints on λ [1], gives

[1] λ must be a conserved quantity for each fluid parcel, friction must be negligible and the fluid must be either barotropic, or λ must be a function of only pressure and density [Pedlosky, 1987, page 39].

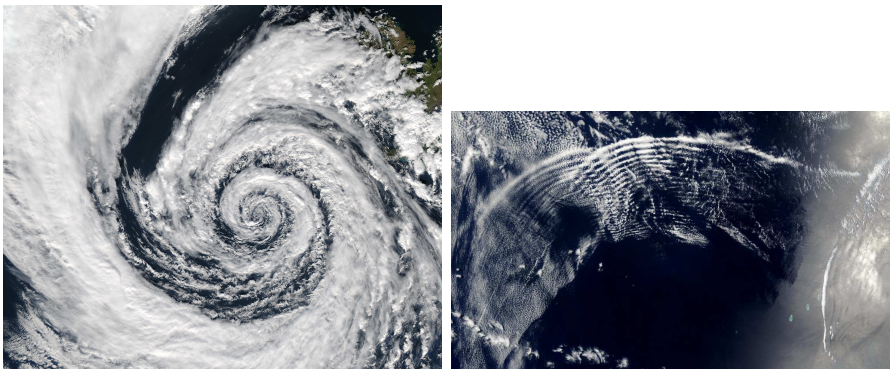


Figure 1.1: Examples of the two different types of motion present in the atmosphere and other geophysical fluids (both images from http://visibleearth.nasa.gov/). On the left is a large-scale low pressure over Iceland. On the right is an example of gravity waves off the coast of Australia. In this case the waves are visible because of the water vapour content of the air.

$$\frac{\mathrm{D}\Pi}{\mathrm{D}t} = 0, \tag{1.2}$$

where $\frac{\mathrm{D}}{\mathrm{D}t}$ indicates the material derivative following the flow. That is, PV is an exactly materially conserved quantity, i.e. the PV is conserved on fluid parcels as they follow the flow. In real geophysical flows this conservation is only approximate but is valid for timescales of around 4 days [McIntyre, 2002b]. The flow can now be described by inversion, or through balance conditions, along with the single evolution equation for PV. These simplified equations can be much faster to solve numerically and provide important insight into the fundamental dynamical process that occur in geophysical flows [see McIntyre, 2002a]. For example, the Rossby wave propagation mechanism can be understood by considering the circulation anomalies induced by perturbing an isoline of PV [see Pedlosky, 1987, p. 102-3]. We will see in section 3.3 that the PV distribution also accounts for the presence of jets in rotating flows.

Although the concept of balance and PV inversion can explain many important large scale motions, there are processes that, by construction, cannot be described within this framework. For example, the breakdown of balance and the subsequent emission of inertia-gravity waves is postulated to be a mechanism whereby energy can be dissipated in the ocean [Molemaker et al., 2005]. Idealised numerical simulations indicate that inertia-gravity waves are an inherent and persistent feature of localised atmospheric jets [Snyder et al., 2007]. Spontaneous emission of inertia-gravity waves from initially balanced flows and their subsequent impact on the balanced flow has been observed in laboratory experiments [Williams et al., 2003] and demonstrated in numerical simulations [Viúdez and Dritschel, 2006].

The question of how much of the fluid motion can be described by the balanced flow is often addressed in the context of the shallow water equations since these are the simplest set of equations that permit inertia-gravity waves. Leith [1980] introduced the concept of a 'slow manifold', that is, a lower dimensional subset of the solution space, devoid of gravity waves, which the flow, if initialised within, will potentially remain in for all time. More recent evidence suggests that such a manifold does not exist. Ford [1994a] considers the evolution of an axisymmetric vortex and demonstrates the existence of a weak instability that, even though too weak to be of practical significance in geophysical fluids, persists even in regimes where the flow could be expected to be balanced. Dritschel and Vanneste [2006] examine the even simpler configuration of a PV front and again see weak instability associated with spontaneous emission of gravity waves. These results indicate that although the balanced manifold does not exist, the effects of the unbalanced motion are frequently weak, even in the parameter regime where balance is expected to break down. In chapter 4 we present more evidence to support this conclusion.

To summarise, the concept of balance is fundamental to our understanding of

geophysical fluid motions which are, in general, close to being balanced. Balanced models reproduce many of the large scale properties of geophysical flows but, due to the non-existence of a true balanced manifold and the associated inevitability of spontaneous emission of inertia-gravity waves, a balanced model will never be sufficient for long term, realistic simulations. This makes the problem of modelling inertia-gravity waves accurately of fundamental importance. Previous studies have quantified the breakdown of balance in the shallow water equations using the most simple configurations of a single vortex [Ford, 1994a] or PV front [Dritschel and Vanneste, 2006]. In this thesis, we look at the most complex (shallow water) configuration: turbulence on a rotating sphere. By understanding the properties of the balanced and unbalanced components of fluid flows and the interaction between them, we can gain a deeper insight into fluid dynamical processes and the limitations of balanced models.

1.2 Geophysical flows

In the previous section we discussed the importance of balance in the context of geophysical flows. In this section we turn our attention to the observed motion of the atmosphere and the common flow features it exhibits.

The motion of the atmosphere and oceans is complex and takes place on a wide range of spatial and temporal scales. Despite the inherent unpredictability of such flows, much can be said about their generic properties due to the presence of coherent structures such as jets (narrow regions of fast flowing fluid) and vortices (areas of spinning fluid). On a rotating planet the zonal background flow, generated by the change in the local projection of the rotation vector with latitude, is perturbed by vortices which interact and merge. The latitudinal variation of rotation also produces a banded structure where regions of homogeneous fluid are bounded by strong zonal jets. These features can easily be seen in images of

the gaseous planets. For example figure 1.2 shows two pictures, one of Jupiter and one of Saturn, that illustrate the banded structure of their atmospheres. Within the bands there are many vortices that are trapped there by the strong zonal jets. These vortices interact and merge with others within their band.

The presence of coherent vortices and jets has important effects on the transport of heat, momentum and chemical or biological tracers [Holloway, 1986]. Vortices effectively sweep these fluid properties along with them while jets act to inhibit cross-jet transport and enhance along-jet transport [Sommeria et al., 1989, Smith, 2005].

In the next section we shall discuss some of the issues with numerically modelling turbulent flows and briefly describe the numerical setup used in this thesis (for more details, see sections 2.5 and 2.6). The following section contains a literature review of the most relevant research into turbulent flows.

1.2.1 Numerical simulation of geophysical flows

Numerical simulation of complex turbulent flows is difficult. The range of spatial and temporal scales poses serious problems for any numerical algorithm. The aim of weather forecasting models is to predict the large scale motions of the atmosphere. Computing power is not yet sufficient that the small scale, fast inertia-gravity waves can be resolved so they are damped by artificial viscous terms and their effects are included via parameterisation schemes. The problem here is that the artificial viscous terms are not selective enough to be benign. Not only do they smear out small scale features, they can also introduce non-trivial spurious dynamics. In this thesis we use a model that has been developed to avoid this problem. By construction, diffusion takes place only at a specified small scale that can be controlled by the modeller (this process is called 'contour surgery' and is described in section 2.6). In addition, the equations are formu-

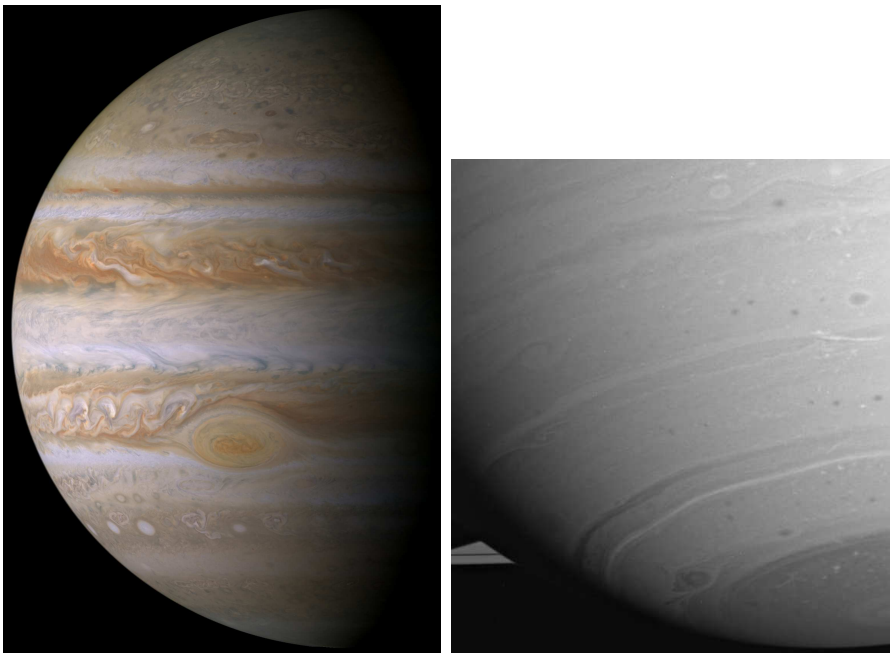


Figure 1.2: Two images of the gaseous planets Jupiter (left, from http://solarsystem.nasa.gov/) and Saturn (right, from http://saturn.jpl.nasa.gov/). Note the zonal banded structure interspersed with small scale coherent vortices. Similar features have been noticed in the atmospheres of other planets and even their moons.

lated in such a way as to make use of the leading order separation between high and low frequency motions. As indicated above this separation is by no means complete but the dominance of the balanced or vortical component in geophysical flows suggests that it is a useful principle. In fact, wave–vortex decomposition techniques are employed in the initialisation of forecast models, the idea being to start the model integration with balanced fields that are as close as possible to the observed state of the atmosphere. However, once the simulations commence, the numerical techniques employed obscure the underlying balance and in modelling the unbalanced component incorrectly, introduce errors into the evolution of the balanced component of the flow. This can be avoided by rewriting the equations so that the underlying structure is apparent. Section 2.5 contains more details and a discussion of the effect this has on the numerics.

In this thesis we shall be solving the shallow water equations. These are the simplest equations that permit both the balanced and unbalanced motions described in section 1.1. However, due to the weakness of the unbalanced motion throughout much of the atmosphere and oceans, it is possible to create even simpler models that share the qualitative behaviour of geophysical flows in certain regimes. The flow regimes can be characterised by the dimensionless parameters summarised in table 1.1. The Froude number, Fr, indicates the importance of gravity and the Rossby number, Ro, indicates the importance of rotation. The squared ratio of Rossby to Froude number is the Burger number, Bu, and it indicates the relative importance of gravity and rotation. The Burger number can also be expressed in terms of the Rossby deformation radius, $L_D = \sqrt{gH}/f$, and a characteristic length scale, L (see table 1.1). The Rossby deformation radius is the length scale at which the effects of gravity and rotation on the deformation of the free surface are approximately equal.

In the limit of small Rossby number, rotation is important. If the Froude number is also small then the shallow water equations reduce to the equivalent

Parameter	**Symbol**	**Definition**
Rossby number	Ro	$\dfrac{U}{fL}$
Froude number	Fr	$\dfrac{U}{\sqrt{gH}}$
Burger number	Bu	$\dfrac{\mathrm{Ro}^2}{\mathrm{Fr}^2}$ or $\dfrac{L_D^2}{L^2}$

Table 1.1: Definitions of the dimensionless parameters that define a flow regime. H, L and U are characteristic depth, length and velocity scales of the flow, f is the Coriolis parameter, g is the acceleration due to gravity, and $L_D = \sqrt{gH}/f$ is the Rossby deformation radius.

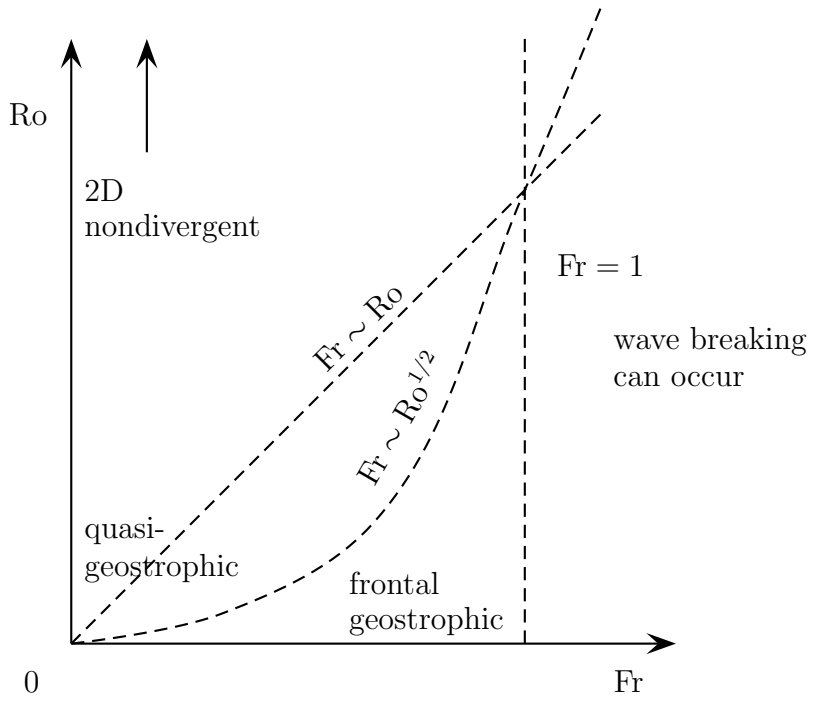


Figure 1.3: Parameter space diagram to illustrate the different limits of the shallow water equations.

$\mathrm{Ro} \ll 1$	$\mathrm{Fr} \ll 1$	equivalent barotropic f plane
$\mathrm{Bu} \gg 1$	$\mathrm{Ro} \ll 1$	2D incompressible (also called barotropic nondivergent f plane)
$\mathrm{Bu} \sim O(1)$	$\mathrm{Ro} \ll 1$	quasi-geostrophic
$\mathrm{Ro} \sim \mathrm{Bu} \ll 1$		frontal geostrophic

Table 1.2: Summary of the different regimes of the shallow water equations.

barotropic equation (conservation of quasi-geostrophic (QG) PV together with the QG balance relations, see section 2.3 for details). For much larger Burger numbers, i.e. large Rossby radius, the fluid becomes approximately two dimensional and nondivergent since the free surface is less able to deform.

1.2.2 Literature review

There have been several studies of turbulence in shallow water flows. More plentiful are studies using the simpler geostrophic or 2D incompressible equations. Due to the arguments outlined in the previous section, these results are still relevant, in some parameter regimes, to the shallow water case. Here we shall give a brief overview of these results.

There are two quite distinct approaches to studying turbulence: spectral and physical. Since Kolmogorov's paper of 1941, it has been common to examine the properties of the energy and enstrophy (half the squared vorticity) spectra. The idea is to consider what happens when the fluid is excited over an initially concentrated range of wavenumbers. Constraints on the conservation of energy and enstrophy in two dimensional turbulence lead to a cascade of enstrophy to smaller scales and an inverse cascade of energy to larger scales [Fjörtoft, 1953]. For three dimensional turbulence the energy cascade is in the reverse direction, i.e. to small scales. Using dimensional arguments, it is possible to derive power laws for the slope of the spectrum in terms of the wavenumber [Salmon, 1998].

Although spectra have been observed that agree with the power laws derived, there are many theoretical reasons to question the applicability of these results (see section 4.7 of Salmon [1998] for a discussion). In particular the presence of coherent structures, ubiquitous in turbulent flows [Fornberg, 1977, McWilliams, 1984, Dritschel, 1993, and many others since], disturbs the assumption of scale similarity upon which the spectral arguments are based [Santangelo et al., 1989]. It is also unclear how to interpret the spectral results in terms of physical space structures [Dritschel, 1993]. For these reasons, in this thesis we shall focus on the physical, rather than the spectral, properties of turbulence.

Given that coherent vortices are a generic feature of turbulent flows, much effort has been focused on observing, analysing and simulating their behaviour. Under the QG approximation vortices of opposite sign behave identically apart from their direction of rotation. However, once the flow departs from geostrophy an asymmetry develops between cyclones and anticyclones. Observational evidence indicates that the asymmetry can be skewed either way depending on the parameter regime. For example, anticyclones are observed to dominate (in terms of a simple population count) the atmospheres of the gas giants [Vasavada and Showman, 2005] and mesoscale oceanic motions [McWilliams, 1985] while cyclones dominate in the tropospheric midlatitudes [Venn, 1887, Wirth, 2000, and references therein], near the extratropical tropopause [Hakim et al., 2002, Hakim and Canavan, 2005] and in small scale oceanic eddies [Eldevik and Dysthe, 2002]. This asymmetry has been investigated using a range of different extensions to the quasi-geostrophic equations and, more recently, using the full shallow water equations.

Before describing the results of these turbulence simulations, we shall outline the information that can be gleaned from considering the properties of individual vortices. First note that anticyclonic vortices are associated with elevations in the free surface height, h, corresponding to high pressure anomalies whereas cyc-

lonic vortices are associated with depressions in h, corresponding to low pressure anomalies. The variations in h associated with the vortex affects the local Rossby deformation radius, $L_D^{loc} \propto \sqrt{gh}$, making it larger in the case of anticyclones and smaller in the case of cyclones. This means that cyclones tend to be smaller and anticyclones more spread out. This is confirmed by observations [Hakim et al., 2002] and numerical simulations [Poulin and Flierl, 2003, Stegner and Dritschel, 2000]. Polvani et al. [1994a] suggest that the local variation in L_D also affects the stability of the vortices since, for vortices of radius larger than the local L_D, the two opposite sides of the vortex cannot 'see' each other and hence cannot act in a coherent fashion.

Although the cyclonic vortices are limited in size, the anticyclonic vortices are limited in strength. Gradient wind balance imposes a limit on the strength of negative pressure gradients (measured from the centre to the outside of the vortex) hence limiting the development of anticyclonic vorticity while imposing no limit on the development of cyclonic vorticity [Olson, 1991, Poulin and Flierl, 2003]. Poulin and Flierl [2003] and Holton [1992] point out that increased anticyclonic vorticity is also curtailed by the criteria for inertial (centrifugal) stability.

Anticyclones and cyclones also differ in their shape. Anticyclones are commonly observed to be axisymmetric whereas cyclones become elongated [Kloosterziel and van Heijst, 1991, Arai and Yamagata, 1994, Stegner and Dritschel, 2000] and can be deformed into triangular or boomerang shapes [Poulin and Flierl, 2003]. Arai and Yamagata [1994] show that this elongation can sometimes lead to cyclones breaking apart; a result that might suggest that cyclones should be more common, even if the numerical simulations do not have sufficient resolution to capture them. A more serious problem with numerical simulations of vortices is highlighted by Dritschel [1998] who points out that the axisymmetrisation of vortices noted in other simulations is likely to be an artifact of numerical dissipation. Naturally occurring vortices are bounded by sharp gradients of vor-

ticity and the failure to represent these in numerical models results in a false picture of vortex shape and stability. As it is the cyclonic vortices that tend to be significantly non-axisymmetric, it is reasonable to question the conclusion that anticyclonic vortices dominate in these simulations.

Cyclone-anticyclone asymmetry is not distinct from the concept of balance discussed in the previous section. Several studies have shown that the amount of imbalance present in the flow can have a profound effect on the vorticity asymmetry [Kuo and Polvani, 2000]. Farge and Sadourny [1989] show that, despite observing no energy transfer between the balanced and unbalanced components of the motions, the two components can indeed affect each other's energy distribution. Ford [1994b] shows that, in some parameter regimes, vortices radiate gravity waves which can cause disruption to the vortex. This is also seen in the results of Polvani et al. [1994a] who observed the emanation of shock like waves from anticyclonic vortices.

As illustrated by the pictures of Jupiter and Saturn (see figure 1.2), in addition to coherent vortices, rotating turbulent flows also exhibit zonal jets. These features are ubiquitous in the atmospheres of the gas giants [Li et al., 2004, Shetty et al., 2007] and have recently been observed in high resolution eddy-permitting simulation of the Earth's oceans [Galperin et al., 2004, Richards et al., 2006].

Jets are important because they influence both the properties of the fluid and the dynamics of the flow. The sharp PV gradients associated with, and responsible for, the jets inhibit the cross-jet transport of chemicals and provide a conduit for Rossby waves. The presence of strong zonal jets can have a profound effect on the dynamics of vortices, confining them within a latitude band or even trapping two vortices together so that they merge (as in the case of Jupiter's white ovals [Youssef and Marcus, 2003]).

Traditional theory [Rhines, 1975] states that jets will form when the Rossby

wave phase speed is comparable to the average velocity of the flow. This theory leads to a way of estimating, via a scale analysis, the number of jets and their separation scale, the Rhines scale, L_{Rh}, based on the fundamental flow parameters, β and U_{jet}, a typical jet velocity. There has been much discussion on how to define U_{jet} and many modifications of L_{Rh} have been suggested [Dritschel and McIntyre, 2008, and references therein]. More recently it has been suggested that a critical latitude exists [Theiss, 2004, Showman, 2007] above which jets will not form. This is due to the latitudinal variation of the Rossby radius, L_D, which appears in the definition of the Rossby wave frequency. At a critical latitude, the Rossby radius, L_D becomes less than the Rhines scale, L_{Rh} and this can be related [Theiss, 2004] to a decrease in anisotropy between the zonal and meridional energy containing scales. Since zonal alternating flows are an extreme example of highly anisotropic flows, this decrease in anisotropy inhibits their formation.

Jet formation can also be thought of in terms of 'PV staircases' [Danilov and Gurarie, 2004, Dritschel and McIntyre, 2008]. PV mixing, caused by a variety of mechanisms, creates regions of homogenised PV separated by regions where the PV has a sharp gradient. Through PV inversion, these sharp gradients are associated with fast eastward flow, i.e. jets. The novel numerical algorithm used in this thesis enables us to simulate jets more accurately than traditional models because the PV gradients on which the jets depend are maintained and not diffused.

1.3 Overview of thesis

In the next chapter we present the mathematical and numerical setup of the shallow water model used in the rest of the thesis. We summarise the relevant properties of the shallow water equations and show how these properties can be exploited to produce a numerical model of unprecedented accuracy and resolution.

Chapter three contains the results of a large number of simulations of shallow water turbulence on a rotating sphere. Initially we focus on four cases from the extreme corners of the Froude number and Rossby number parameter space. We examine cyclone–anticyclone asymmetry and jet formation in detail for these cases and then generalise our results across the parameter space.

In chapter four we return to the subject of balance. Using four different definitions of the balanced component of a flow, we quantify the extent to which the turbulent flows of chapter three remain balanced throughout their evolution. In addition we attempt to quantify the impact of the unbalanced component of the flow on the overall motion.

Chapter five consists of our conclusions and some suggestions for further work.

Chapter 2

Methods

In this chapter we outline the governing equations for fluid motion and provide a brief explanation of how they can be simplified under the assumption that the fluid layer is shallow (i.e. the horizontal scale of the motion is much larger than the vertical scale). In section 2.2 we describe the frequency separation that distinguishes inertia-gravity waves from Rossby waves in the linear equations about a state of rest. We then outline the relationship between the shallow water system and the simpler quasi-geostrophic model. Then follows a description of some of the properties of the shallow water system with particular focus on those relevant to the setup of the model used in this thesis. This model, the Contour Advective Semi-Lagrangian algorithm, is described in section 2.6 along with some numerical details.

2.1 Governing equations

The Navier Stokes equations that govern the motion of a homogeneous, incompressible fluid are well known and have been extensively studied. Due to their complexity, many simplifications based on scaling arguments have been de-

rived. These are well documented and details may be found in textbooks such as Gill [1982] and Pedlosky [1987]. Those approximations relevant to this study are briefly outlined below.

The Navier Stokes equations comprise three momentum equations and the continuity equation. For the homogeneous, incompressible fluid considered here, the continuity equation is simply

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0, \tag{2.1}$$

where $\boldsymbol{u}$ is the fluid velocity.

In a coordinate system rotating with constant angular velocity $\boldsymbol{\Omega}$, the momentum equations can be written in vector form as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\frac{1}{\rho}\nabla p + \nabla\Phi - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) + \mathscr{F}, \tag{2.2}$$

where $\mathrm{D}/\mathrm{D}t \equiv \frac{\partial}{\partial t} + \boldsymbol{u}.\nabla$ is the material derivative, ρ the density, p the pressure and $\nabla\Phi$ and $\mathscr{F}$ represent the conservative and non-conservative forces respectively. Two of the terms in equation 2.2 arise purely because of the rotation of the coordinate system. The Coriolis acceleration, $2\boldsymbol{\Omega} \times \boldsymbol{u}$, acts to curve fluid parcel trajectories to the right (left) in the Northern (Southern) hemisphere. The centripetal acceleration, $\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r})$, can be written as the gradient of a potential and incorporated into the definition of Φ. Φ is then called the geopotential. The sphere is, to good approximation, a surface of constant Φ so in this case $\nabla\Phi$ becomes (0, 0, -g), the acceleration due to gravity. For the purposes of this study, non-conservative forces will be neglected, i.e. $\mathscr{F} = 0$. We will be modelling these equations for a shallow fluid layer on a sphere. However, spherical geometry complicates the theoretical properties of the equations and it is easier to work in local cartesian coordinates. See White [2002] for the full spherical shallow water

equations and a brief discussion of the effects of spherical geometry.

Writing out the components of equation 2.2 in local cartesian coordinates and including the simplifications outlined above gives

$$\frac{\mathrm{D}u}{\mathrm{D}t} + 2\Omega(w\cos\phi - v\sin\phi) = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{2.3}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + 2\Omega u\sin\phi = -\frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{2.4}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} - 2\Omega u\cos\phi = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g. \tag{2.5}$$

If we now consider the vertical component for a fluid at rest we obtain

$$\frac{1}{\rho}\frac{\partial p}{\partial z} = -g. \tag{2.6}$$

This equation states that the vertical pressure gradient is balanced by gravity. This is known as hydrostatic balance. Using the shallow water scaling, we can show that it is consistent to assume that equation 2.6 holds even when the fluid is in motion [Pedlosky, 1987]. This is the hydrostatic approximation and is sometimes used as the definition of the shallow water model [Pedlosky, 1987]. In taking the hydrostatic approximation we assume that the vertical acceleration is negligible compared to the vertical pressure gradient. We also assume that the term $2\Omega u\cos\phi$ due to the horizontal variation of Ω is negligible. Note that to conserve energy, the corresponding $2\Omega w\cos\phi$ term in equation 2.3 must also be neglected [Salmon, 1998, White, 2002].

Integrating equation 2.6 with the boundary condition that $p(x, y, h) = p_0$ gives

$$p(z) = \rho g(h - z) + p_0. \tag{2.7}$$

Using equation 2.7 to eliminate p from 2.3 and 2.4 we have

$$\frac{\mathrm{D}u}{\mathrm{D}t} - fv = -g\frac{\partial h}{\partial x}, \tag{2.8}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + fu = -g\frac{\partial h}{\partial y}, \tag{2.9}$$

where $f = 2\Omega \sin\phi$ is the Coriolis parameter. Equations 2.8-2.9 show that if the horizontal velocities u and v are initially independent of depth they will remain so as the Coriolis and pressure gradient forces are also independent of depth. This enables us to integrate equation 2.1 from $z = 0$ (the lower boundary) to $z = h$ (the depth of the fluid) to obtain

$$h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + w(h) - w(0) = 0. \tag{2.10}$$

At $z = 0$ there can be no flow normal to the boundary so $w(0) = 0$ while at the free surface $w(h) = Dh/Dt$. So, in vector form, equations 2.8 - 2.10 can be written

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + f\boldsymbol{k} \times \boldsymbol{u} = -g\boldsymbol{\nabla} h \tag{2.11}$$

$$\frac{\partial h}{\partial t} + \boldsymbol{\nabla} \cdot (h\boldsymbol{u}) = 0 \tag{2.12}$$

where $\boldsymbol{u}$ is the horizontal velocity and only the two components perpendicular to $\boldsymbol{k}$, the local vertical, of equation 2.11 are considered.

The shallow water equations 2.11-2.12 are commonly expressed in terms of the vertical component of the relative vorticity $\zeta = v_x - u_y$ and the divergence $\delta = u_x + v_y$:

$$\frac{\mathrm{D}\zeta}{\mathrm{D}t} + (\zeta + f)\delta - \beta v = 0, \tag{2.13}$$

$$\frac{\mathrm{D}\delta}{\mathrm{D}t} + \delta^2 - 2J(u, v) - f\zeta + \beta u = -g\nabla^2 h, \tag{2.14}$$

where $\beta = \partial f/\partial y$. Rewriting equation 2.12 in the form

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\delta = 0 \tag{2.15}$$

enables us to eliminate δ from equation 2.13 to give

$$\frac{D}{Dt}\left(\frac{\zeta + f}{h}\right) \equiv \frac{\mathrm{D}\Pi}{\mathrm{D}t} = 0, \tag{2.16}$$

where

$$\Pi \equiv \frac{\zeta + f}{h} \tag{2.17}$$

is the potential vorticity (PV). Equation 2.16 states that the PV is conserved following a fluid parcel. Although equation 2.16 has been derived using several assumptions, in particular the absence of forcing, it is still approximately true in atmospheric and oceanic flows provided the timescale of the flow under consideration is short enough ($<$ 4 days McIntyre [2002b]).

2.2 Linearised equations

In chapter 1 we stated that the PV controls the dominant large scale 'balanced' motion. Here we show this using the linear equations.

Linearising equations 2.11-2.12 about a state of rest ($\boldsymbol{u} = 0, h = H$) and taking f to be constant gives

$$\frac{\partial u'}{\partial t} - fv' = -g\frac{\partial h'}{\partial x}, \tag{2.18}$$

$$\frac{\partial v'}{\partial t} + fu' = -g\frac{\partial h'}{\partial y}, \tag{2.19}$$

$$\frac{\partial h'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0, \tag{2.20}$$

where the primed variables represent small departures from the background state. Assuming solutions of the form

$$q' = \hat{q}e^{i(kx+ly-\omega t)} \tag{2.21}$$

(real part intended) for each variable gives

$$-i\omega\hat{u} - f\hat{v} = -ikg\hat{h}, \tag{2.22}$$

$$-i\omega\hat{v} + f\hat{u} = -ilg\hat{h}, \tag{2.23}$$

$$-i\omega\hat{h} + iH(k\hat{u} + l\hat{v}) = 0. \tag{2.24}$$

Equations 2.22-2.24 only have non-zero solutions if the dispersion relation is satisfied, i.e.

$$\omega(\omega^2 + f^2 - K^2gH) = 0, \tag{2.25}$$

where $\mathbf{K} = (k, l)$ is the horizontal wavenumber vector so $K^2 = k^2 + l^2$. Equation 2.25 has solutions

$$\omega = 0, \pm\sqrt{f^2 + K^2gH}. \tag{2.26}$$

The solutions $\omega = \pm\sqrt{f^2 + K^2gH}$ correspond to inertia-gravity waves. Note that

$$|\omega_{IG}| = |\sqrt{f^2 + K^2 gH}| \geq f \tag{2.27}$$

so that there is a distinct separation in frequency between these solutions and the $\omega = 0$ solution. For 'short' wavelengths $\omega \sim K\sqrt{gH}$ which is the dispersion relation for gravity waves in the absence of rotation. For 'long' wavelengths $\omega \sim f$ which gives rise to inertial motion. The length scale that determines whether the wavelength is short or long is

$$L_D = \frac{\sqrt{gH}}{f}. \tag{2.28}$$

This is the Rossby deformation radius and it is the length scale above which rotation becomes important.

Another property of the inertia-gravity wave solutions is that they have zero perturbation potential vorticity. The perturbation PV is found by linearising the PV about $\zeta = 0, h = H$ and is defined as

$$\Pi' = \frac{\zeta'}{H} - \frac{fh'}{H^2}. \tag{2.29}$$

The linearised version of equation 2.16 is

$$H\frac{\partial \zeta'}{\partial t} - f\frac{\partial h'}{\partial t} = 0. \tag{2.30}$$

Assuming solutions of the form 2.21 we have

$$-i\omega(H\hat{\zeta} - f\hat{h}) = 0, \tag{2.31}$$

which implies that for non zero ω, $H\hat{\zeta} - f\hat{h} = 0$ or equivalently, $\Pi' = 0$.

We now consider the other solution, $\omega = 0$, to equation 2.25. Substituting $\omega = 0$ into equations 2.22-2.23 gives

$$-f\hat{v} = -ikg\hat{h}, \tag{2.32}$$

$$f\hat{u} = -ilg\hat{h}. \tag{2.33}$$

from which we can calculate that

$$\delta = 0, \tag{2.34}$$

so that these waves are non-divergent. Now linearise equation 2.14 (with constant f i.e. $\beta = 0$) to obtain

$$\frac{\partial \delta}{\partial t} - f\zeta' = -g\nabla^2 h' \tag{2.35}$$

and again assume solutions of the form 2.21 for each variable. With $\omega = 0$ this gives

$$f\hat{\zeta}' - K^2 g\hat{h}' = 0, \tag{2.36}$$

so that for these waves

$$\gamma' = f\zeta' - g\nabla^2 h' = 0. \tag{2.37}$$

Equation 2.37 is a form of the geostrophic equations

$$fv' = g\frac{\partial h'}{\partial x}, \tag{2.38}$$

$$-fu' = g\frac{\partial h'}{\partial y}. \tag{2.39}$$

These equations are diagnostic balance relations which state that the Coriolis acceleration is balanced by the pressure gradient. They are valid when the acceleration term in equation 2.11 can be neglected. Taking U and L to be typical

horizontal velocity and length scales respectively the acceleration and Coriolis terms scale like

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} \sim \frac{U^2}{L}, \tag{2.40}$$

$$f \times \boldsymbol{u} \sim fU, \tag{2.41}$$

so the condition that the acceleration term is negligible compared to the Coriolis term is equivalent to

$$\mathrm{Ro} \equiv \frac{U}{fL} \ll 1. \tag{2.42}$$

Ro is the Rossby number, a dimensionless number that measures the significance of rotation in the flow. For typical geophysical flows, the Rossby number is small compared to unity (Ro is roughly 0.1-0.5 in the atmosphere and 0.01-0.2 in the ocean).

On the sphere f is not constant but varies with latitude. We can include the first order approximation to this by setting $f = f_0 + \beta y$. Under this assumption the linearised shallow water equations 2.18-2.20 no longer have constant coefficients so the dispersion relation cannot be derived in the same way. However we can examine the geostrophic mode, i.e. the mode governed by the PV, using the linearised form of equation 2.16:

$$H\frac{\partial \zeta'}{\partial t} - f_0\frac{\partial h'}{\partial t} + H\beta v = 0. \tag{2.43}$$

Substituting for ζ and v using the geostrophic relations (equations 2.37 and 2.39) with f replaced by f_0, we have an equation entirely in terms of one variable h':

$$\frac{\partial \nabla^2 h'}{\partial t} - L_D^{-2}\frac{\partial h'}{\partial t} + \beta\frac{\partial h'}{\partial x} = 0. \tag{2.44}$$

Assuming solutions of the form 2.21 as before yields the dispersion relation

$$\omega = -\frac{k\beta}{k^2 + l^2 + L_D^{-2}}. \tag{2.45}$$

This is the Rossby wave solution. When $\beta = 0$ it corresponds to the $\omega = 0$ solution. The other two roots do not appear because the gravity waves are not permitted by equation 2.44.

To summarise, we have demonstrated that the linear shallow water equations permit two distinct types of solution corresponding to two distinct types of motion: the zero (or low) frequency Rossby wave and the high ($> f$) frequency inertia-gravity waves. These two types of motion are fundamentally different.

2.3 A prognostic equation for the slow modes

In the previous section we saw that if the Rossby number is small the acceleration term in equation 2.11 can be neglected, providing us with two diagnostic equations for u and v given h. This can be more easily seen if we non-dimensionalise equations 2.11-2.12. To do this we choose appropriate length and velocity scales L and U that characterise the flow:

$$(x, y) = L(x', y'), \tag{2.46}$$

$$u = Uu', \tag{2.47}$$

where the primes denote dimensionless variables. We assume that the relevant time scale for the flow is $T = L/U$, i.e. the advective rather than the gravity wave time scale. Substituting this into equation 2.11 and applying the beta-plane approximation $f = f_0 + \beta y$ gives

$$\frac{U^2}{L}\frac{\mathrm{D}\boldsymbol{u}'}{\mathrm{D}t'} + (f_0 + \beta L y')U\boldsymbol{k}\times\boldsymbol{u}' = -\frac{gD}{L}\boldsymbol{\nabla}'\tilde{h} \tag{2.48}$$

$$\frac{UD}{L}\frac{\partial\tilde{h}}{\partial t'} + \frac{UD}{L}\boldsymbol{\nabla}'\cdot(\tilde{h}\boldsymbol{u}') + \frac{UH}{L}\boldsymbol{\nabla}'\cdot\boldsymbol{u}' = 0 \tag{2.49}$$

where $\tilde{h} = (h-H)/H$ is the nondimensional depth anomaly and D is the scale of variation in free surface height, $h - H$. Since we are interested in the case where $\mathrm{Ro} = U/f_0L < 1$, we require, as in the previous section, the pressure gradient terms to balance, to leading order, the Coriolis terms. From equation 2.48, this gives a scaling for D:

$$D \sim \frac{f_0 UL}{g}. \tag{2.50}$$

Now equations 2.48 and 2.49 can be written in nondimensional form as

$$\frac{U}{f_0 L}\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \left(1 + \frac{\beta L y}{f_0}\right)\boldsymbol{k}\times\boldsymbol{u} = -\boldsymbol{\nabla}\tilde{h}, \tag{2.51}$$

$$\frac{U}{f_0 L}\left[\frac{\partial\tilde{h}}{\partial t} + \boldsymbol{\nabla}\cdot(\boldsymbol{u}\tilde{h})\right] + \frac{gH}{f_0^2 L^2}\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{2.52}$$

where the primes have been dropped for clarity. We recognise the dimensionless parameter U/f_0L as the Rossby number. The remaining dimensionless parameter is the Burger number Bu and it can be expressed in several ways:

$$\mathrm{Bu} = \frac{gH}{f_0^2 L^2}, \tag{2.53}$$

$$= \frac{L_D^2}{L^2}, \tag{2.54}$$

$$= \frac{\mathrm{Ro}^2}{\mathrm{Fr}^2}, \tag{2.55}$$

where L_D is the Rossby deformation radius as defined in equation 2.28 and

$$\mathrm{Fr} = \frac{U}{\sqrt{gH}} \tag{2.56}$$

i.e. the ratio of the flow speed to the gravity wave speed, is the Froude number. In terms of these parameters, equations 2.51 and 2.52 become

$$\mathrm{Ro}\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + (1+\hat{\beta}y)\boldsymbol{k}\times\boldsymbol{u} = -\boldsymbol{\nabla}\tilde{h}, \tag{2.57}$$

$$\mathrm{Ro}\left[\frac{\partial\tilde{h}}{\partial t} + \boldsymbol{\nabla}\cdot(\boldsymbol{u}\tilde{h})\right] + \mathrm{Bu}\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0. \tag{2.58}$$

where $\hat{\beta} = \beta L/f_0$. We now assume that the variation in f is small compared to the horizontal scale of the flow, i.e. $\hat{\beta} \ll 1$. Taking Bu $\sim O(1)$, $\hat{\beta} = \tilde{\beta}\mathrm{Ro}$ and expanding each variable q as an asymptotic series in Ro, $q = q_0 + \mathrm{Ro}\, q_1 + ...$, we can compare powers of Ro in equations 2.57-2.58. The $O(1)$ terms in equations 2.57-2.58 are the geostrophic balance relations

$$\boldsymbol{k}\times\boldsymbol{u}_0 = -\boldsymbol{\nabla}\tilde{h}_0, \tag{2.59}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}_0 = 0. \tag{2.60}$$

Looking at the $O(\mathrm{Ro})$ terms gives

$$\frac{\partial\boldsymbol{u}_0}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\boldsymbol{u}_0 + \tilde{\beta}y(\boldsymbol{k}\times\boldsymbol{u}_0) + \boldsymbol{k}\times\boldsymbol{u}_1 = -\boldsymbol{\nabla}\tilde{h}_1, \tag{2.61}$$

$$\frac{\partial\tilde{h}_0}{\partial t} + \boldsymbol{u}_0\cdot\boldsymbol{\nabla}\tilde{h}_0 + \boldsymbol{\nabla}\cdot\boldsymbol{u}_1 = 0, \tag{2.62}$$

which can be rearranged to give a single prognostic equation in $\tilde{h}_0$:

$$\frac{\partial}{\partial t}\left(\nabla^2\tilde{h}_0 - \tilde{h}_0\right) + \frac{\partial\tilde{h}_0}{\partial x}\frac{\partial}{\partial y}\nabla^2\tilde{h}_0 + \frac{\partial\tilde{h}_0}{\partial x}\tilde{\beta} - \frac{\partial\tilde{h}_0}{\partial y}\frac{\partial}{\partial x}\nabla^2\tilde{h}_0 = 0. \tag{2.63}$$

This equation is equation 2.16 for the non-dimensional quasi-geostrophic potential vorticity. To see this, we first rewrite the PV definition, equation 2.17, in terms of scaled variables

$$\Pi = \frac{\frac{U}{L}\zeta' + f_0(1+\hat{\beta}y')}{H(1+\frac{D}{H}\tilde{h})} \tag{2.64}$$

$$= \frac{f_0}{H}\frac{\mathrm{Ro}\zeta' + 1 + \mathrm{Ro}\tilde{\beta}y'}{1+\mathrm{RoBu}^{-1}\tilde{h}'}. \tag{2.65}$$

Since we are assuming $\mathrm{Bu} \sim O(1)$ and $\mathrm{Ro} \ll 1$, this becomes

$$\Pi \approx \frac{f_0}{H}(\mathrm{Ro}\zeta' + 1 + \mathrm{Ro}\tilde{\beta}y')(1-\mathrm{Ro}\tilde{h}'), \tag{2.66}$$

$$\approx \frac{f_0}{H}(\mathrm{Ro}\zeta' + 1 + \mathrm{Ro}\tilde{\beta}y' - \mathrm{Ro}\tilde{h}'). \tag{2.67}$$

$$\tag{2.68}$$

Dropping the primes and expanding variables as series in Ro as before gives

$$\Pi \approx \frac{f_0}{H}(1+\mathrm{Ro}(\zeta_0 - \tilde{h}_0 + \tilde{\beta}y)), \tag{2.69}$$

$$= \frac{f_0}{H}(1+\mathrm{Ro}(\nabla^2\tilde{h}_0 - \tilde{h}_0 + \tilde{\beta}y)), \tag{2.70}$$

where we have used the geostrophic equation 2.59 to replace ζ_0 with $\nabla^2\tilde{h}_0$. Dropping the constant term and factor, which has no effect on the evolution equation, we have the quasi-geostrophic potential vorticity:

$$\Pi_{QG} = \nabla^2\tilde{h}_0 - \tilde{h}_0 + \tilde{\beta}y. \tag{2.71}$$

Substituting this into equation 2.16 and using the geostrophic velocity as the advecting velocity yields equation 2.63.

Equation 2.63 and the balance relations 2.59 make up the (equivalent barotropic) quasi-geostrophic model - the most common geophysical balanced model. It contains just one prognostic equation and does not permit gravity wave solutions. This makes it simpler to study than the full shallow water equations.

2.4 Equatorial dynamics

The geostrophic scaling presented in the previous section is valid over much of the sphere but breaks down towards the equator as $f \to 0$. Following Wheeler [2002] we take the equatorial β plane approximation, $f = \beta y$. Equations 2.18-2.20 become

$$\frac{\partial u'}{\partial t} - \beta y v' = -g\frac{\partial h'}{\partial x}, \tag{2.72}$$

$$\frac{\partial v'}{\partial t} + \beta y u' = -g\frac{\partial h'}{\partial y}, \tag{2.73}$$

$$\frac{\partial h'}{\partial t} + H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0. \tag{2.74}$$

Following a similar procedure as in section 2.2, we look for wave solutions of the form

$$q = \hat{q}(y)e^{i(kx-\omega t)}. \tag{2.75}$$

Substituting this into equations 2.72-2.74 and rearranging gives an equation in $\hat{v}$:

$$\frac{\partial^2 \hat{v}}{\partial y^2} + \left(\frac{\omega^2}{gH} - k^2 - \frac{k\beta}{\omega} - \frac{\beta^2 y^2}{gH}\right)\hat{v} = 0, \tag{2.76}$$

whose solutions are given in terms of parabolic cylinder functions. In order for the β plane approximation to remain valid, solutions of this equation must decay away from $y = 0$. This happens if and only if

$$\frac{\sqrt{gH}}{\beta}\left(\frac{\omega^2}{gH} - k^2 - \frac{k\beta}{\omega}\right) = 2n+1, \qquad n = 0, 1, 2, \ldots. \tag{2.77}$$

For $n \neq 0$ this yields similar solutions as in section 2.2, i.e. two roots correspond to inertia-gravity waves and one root corresponds to a Rossby wave. For $n = 0$ we obtain the mixed Rossby-gravity, or Yanai, wave.

Another special solution of equations 2.72-2.74 can be obtained by setting $\hat{v} = 0$. This gives the dispersion relation

$$\omega = \sqrt{gH}k, \tag{2.78}$$

which describes the equatorial Kelvin wave. The meridional structure of this wave is given by

$$\hat{u}(y) \propto \exp\left(\frac{-\beta y^2}{\sqrt{gH}}\right). \tag{2.79}$$

The dispersion relations 2.77 and 2.78 are plotted in figure 2.1.

2.5 The shallow water equations transformed

The shallow water equations are the simplest set of equations to describe both balanced motion and inertia-gravity waves. As such they provide an appropriate test for both new numerical algorithms and wave–vortex decomposition techniques.

The shallow water equations as written above have been derived by considering the various physical forces acting on fluid parcels. They are written in terms of the variables (layer depth h and horizontal velocity $\boldsymbol{u}$) that are directly useful for visualising the fluid flow. However, this form of the equations hides some underlying mathematical properties such as the Lagrangian conservation of PV

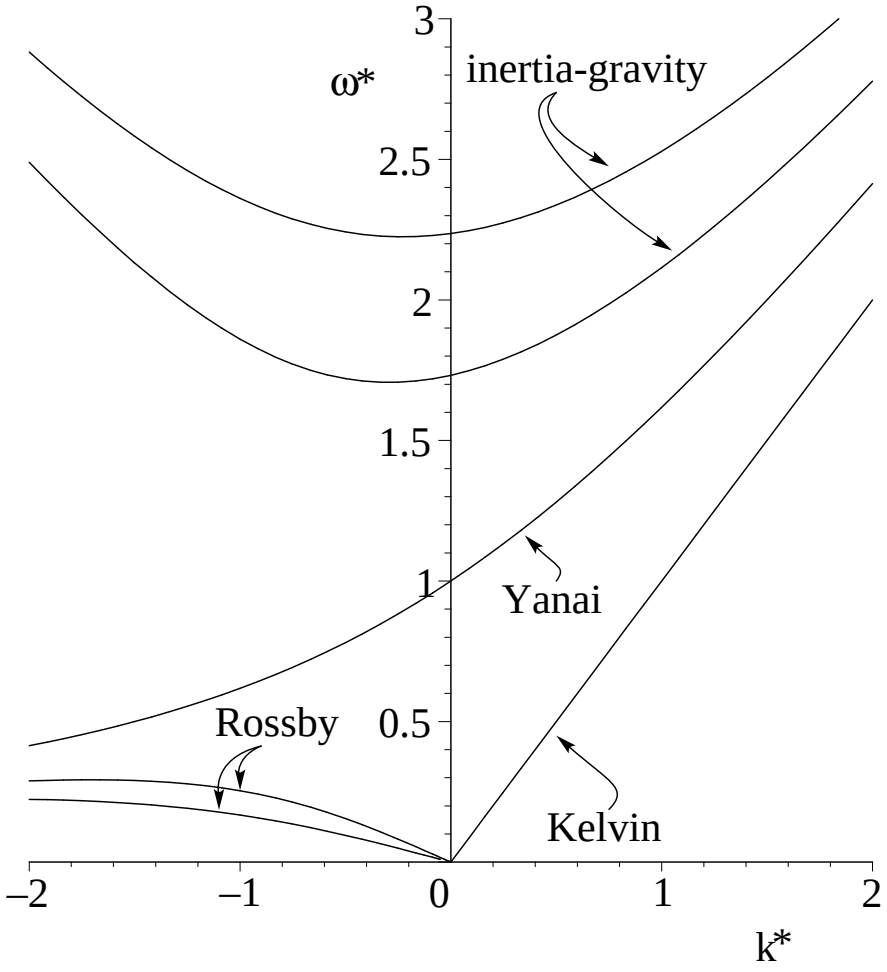


Figure 2.1: Dispersion curves for the equatorial waves plotted in terms of nondimensional frequency $\omega^* = \omega/(\beta\sqrt{gH})^{1/2}$ and wavenumber $k^* = k(\sqrt{gH}/\beta)^{1/2}$.

(equation 2.16). It has been shown [Dritschel and Viúdez, 2003, Mohebalhojeh and Dritschel, 2000b] to be beneficial to rewrite the equations in terms of different variables. Since the PV field is central to the balanced dynamics *and* has the property that it is materially conserved, Π is a natural choice for one of the prognostic variables. This is unconventional but not unprecedented. Thuburn [1997], Bates et al. [1995] and Temperton and Staniforth [1987] for example, each employ models that use PV as a prognostic variable. Here we use the contour advection method developed by Dritschel and Ambaum [1997] and Dritschel et al. [1999]. This method is outlined in the following section.

Given equation 2.16 as one of our prognostic equations, we now have to choose the other two variables. A common variable choice is $(\zeta, \delta, \tilde{h})$ so an obvious choice would be to replace the vorticity ζ with potential vorticity Π and keep the other two variables. However, small errors in calculating the nonlinear terms in the prognostic equation for $\tilde{h}$ can result in erroneous gravity waves in the divergence field [Dritschel and Mohebalhojeh, 2000]. Numerous studies (for example Mohebalhojeh and Dritschel [2000b], Dritschel and Viúdez [2003] and Smith and Dritschel [2006]) have shown that it is advantageous to avoid evolving h and instead use prognostic equations for variables that vanish in the limit of vanishing Rossby and Froude numbers. Equations of this form then have the property that geostrophic balance is recovered, as expected, when $\mathrm{Fr} \sim \mathrm{Ro} \ll 1$. Mohebalhojeh and Dritschel [2000a] derive hierarchies of variables for which this is true, in the discrete model. Their hierarchies are based on the variables δ and γ (and their time derivatives), where

$$\delta = \nabla \cdot \boldsymbol{u}, \tag{2.80}$$

is the divergence (as introduced in section 2.1) and

$$\gamma = \nabla \cdot \boldsymbol{a} \tag{2.81}$$
$$= \nabla \cdot \frac{D\boldsymbol{u}}{Dt} \tag{2.82}$$
$$= f\zeta - \beta u - gH\nabla^2\tilde{h}, \tag{2.83}$$

is the divergence of the acceleration $\boldsymbol{a}$. Note that for constant background rotation ($\beta = 0$), γ/f is the ageostrophic vorticity. It has been shown that in this case, i.e. on an f-plane, that δ and γ themselves are the best choice of variables over a wide range of Fr and Ro [Mohebalhojeh and Dritschel, 2000b]. It is not obvious that these variables should work so well on the sphere due to the breakdown of geostrophic balance at the equator. However, we shall see in chapter 4 that they are remarkably effective.

The prognostic equations for δ and γ in spherical coordinates [Smith and Dritschel, 2006][1] are

$$\delta_t - \gamma = -|\boldsymbol{u}|^2 - 2[u_\phi(u_\phi + \zeta) + v_\phi(v_\phi - \delta)] - \boldsymbol{\nabla} \cdot (\delta \boldsymbol{u}), \tag{2.84}$$
$$\gamma_t - c^2\nabla^2\delta = c^2\nabla^2\{\boldsymbol{\nabla} \cdot [(\tilde{h})\boldsymbol{u}]\} + 2\Omega B_\lambda - \boldsymbol{\nabla} \cdot (Z\boldsymbol{u}), \tag{2.85}$$

where $c^2 \equiv gH$ (mean-square gravity wave speed), $B \equiv c^2\tilde{h} + \frac{1}{2}|\boldsymbol{u}|^2$ (Bernoulli pressure), and $Z = f(\zeta + f)$. We now have three prognostic equations for the variables (Π,δ,γ). However, we still need $\tilde{h}$, u and v to solve these equations. Writing $\boldsymbol{u}$ as the sum of a streamfunction and divergence potential we have

$$\boldsymbol{u} = \boldsymbol{k} \times \boldsymbol{\nabla}\psi + \boldsymbol{\nabla}\chi, \tag{2.86}$$

where

[1] some of the work here was completed before publication in Smith and Dritschel [2006]

$$\nabla^2\psi = \zeta \tag{2.87}$$

$$\nabla^2\chi = \delta. \tag{2.88}$$

Equation 2.88 can be solved directly as δ is a prognostic variable. To solve 2.87 we need to find ζ which appears in the definitions of both Π and γ

$$\zeta = (1+\tilde{h})\Pi - f, \tag{2.89}$$

$$\gamma = f\zeta - \beta u - c^2\nabla^2\tilde{h}, \tag{2.90}$$

where $\beta = \partial f/\partial\phi = 2\Omega\cos\phi$ and $\tilde{h} = (h-H)/H$ is the dimensionless depth anomaly. It is possible to eliminate ζ from these equations to obtain an equation for $\tilde{h}$ but this equation will still contain u which depends on ψ (and hence ζ) by equation 2.86. In practise we find ζ from equation 2.89 using a first guess for $\tilde{h}$. We then invert equation 2.87 to obtain $\boldsymbol{u}$ which in turn gives a better guess for $\tilde{h}$ via equation 2.90. This new value of $\tilde{h}$ is plugged into equation 2.89 and the process is iterated until it converges (typically less than 10 iterations for Fr < 0.5, slowly increasing for higher Froude numbers).

2.6 The CASL algorithm

We have described the central importance of the PV to geophysical flows in chapter 1. In order to be able to model the dominant balanced component of the flow accurately it is essential for us to have an accurate PV field. This is made difficult by the tendency for the PV field to develop fine scales such as filaments and sharp gradients. These features are ubiquitous and have a large impact on the evolution of the flow. Sharp PV gradients give rise to jet streams of fast fluid flow

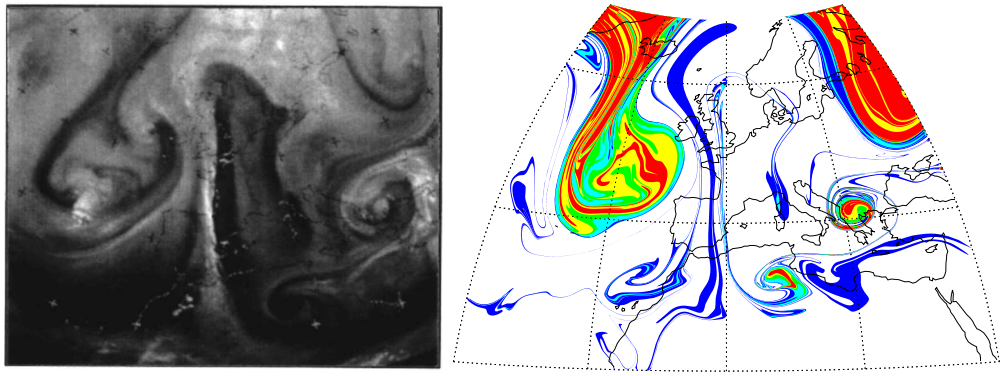


Figure 2.2: Meteosat water vapour image (left) and the corresponding PV field calculated from ECMWF analysis fields using a diagnostic contour advection technique. Taken from Appenzeller et al. [1996].

[Hoskins et al., 1985, Marcus and Lee, 1998, Dritschel and McIntyre, 2008]. These are commonly seen in the atmospheres of the gaseous planets (see figure 1.2) and are also present in the Earth's atmosphere and oceans. Fine scale filaments of PV can roll up into vortices and so influence the large scale weather patterns [Appenzeller et al., 1996]. Figure 2.2 shows two figures taken from Appenzeller et al. [1996]. The first is a water vapour image which shows fine scale structure and filaments. Water vapour images can be used distinguish between air masses with different PV values as these airmasses also have a distinctive water vapour content. The second figure shows the PV field obtained from ECMWF analysis fields using a dynamic contour advection technique [Norton, 1994, Waugh and Plumb, 1994]. The similarity between the two figures is clear, as is the fine scale structure in the PV field. Studies have shown that these fine scale features can have an impact on atmospheric chemistry [McIntyre, 1995].

Fine scale PV structure is not well captured by traditional grid point or spectral methods [Dritschel et al., 1999]. This has motivated a Lagrangian approach [Dritschel and Ambaum, 1997, Dritschel et al., 1999] whereby PV is explicitly materially conserved. The PV field is discretised into levels separated by con-

tours that represent a jump $\Delta\Pi$ in Π. These contours are represented by a set of nodes that are advected by the flow (the nodal positions are updated using an Adams-Bashforth scheme). All other variables are held on a latitude-longitude grid with an equal number of points in each direction [Smith and Dritschel, 2006]. Note that this is different from the usual spherical spectral model grid which has twice the number of points in longitude than in the latitude. The reason for this difference is explained in section 2.7. The flexibility of the contour representation is illustrated in figure 2.3. A single PV contour is shown superimposed on the underlying grid. The nodes are concentrated where the resolution is most needed, i.e. where the contour is most curved. Note that the filament is resolved despite having a width much less than the grid length. The filamentary structure generated by the flow is retained down to a length scale d_s beneath which it becomes necessary to perform a procedure to regularise the contours. This is done in two stages. Firstly the contour surgery routine searches for nodes on contours of the same PV level which are closer than the surgery scale d_s (measured perpendicular to the contour). Nodes that satisfy this condition are then reconnected. Once this has been completed for all nodes the second stage, node redistribution, takes place. This is necessary to ensure that the nodes are representing the contours accurately. The density of the nodes along a contour depends upon the curvature of the contour: a sharp curvature will require more nodes. Nodes can also be removed from contours where they are no longer required. Typically the PV field is resolved down to a scale one tenth of the latitudinal grid length before surgery is performed to remove the dynamically inactive filaments.

2.7 Numerical setup

The CASL algorithm is unconventional in the way that PV is treated. For the other variables it exploits the efficiency of already established numerical meth-

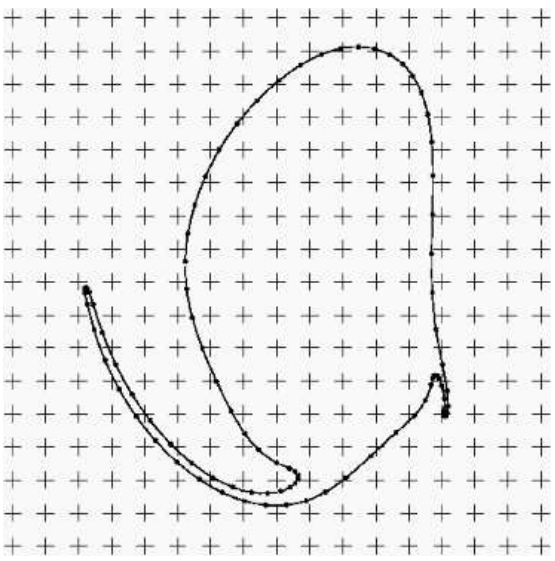


Figure 2.3: An example of a PV contour overlying the grid (crosses) on which the other variables are held. Note the uneven distribution of nodes along the contour: more are needed to represent the highly curved sections.

ods. The solution procedure has been outlined in section 2.5. The method used is semi-spectral, i.e. Fourier series are used in longitude and 2nd order finite differencing in latitude. This requires extra resolution in the latitudinal direction to compensate for the higher errors involved in finite differencing compared to Fourier series. However, as we shall show in section 2.7.2, the formal accuracy of the spectral approach is no greater than that of the 4th order, or even a 2nd order, finite difference approach when the PV field is discontinuous.

The time-stepping procedure for the gridded fields is a standard semi-implicit leapfrog scheme [Ritchie, 1988]. This scheme decouples the even and odd time levels so requires the addition of a filter for stability [Robert, 1966, Asselin, 1972]. The effect of this filter and the appropriate choice of the filter coefficient is discussed in the next section.

Apart from contour surgery and Robert-Asselin time filtering, the only other

numerical stabilization required in the CASL model is the 'Broutman' spectral filter that is applied to the nonlinear tendencies of δ and γ. This filter is only applied in longitude and is required to prevent the build up of grid-scale noise due to aliasing errors.

2.7.1 Robert-Asselin filter coefficient

The Robert-Asselin filter replaces a field q at time t with a combination of the field at $t - \Delta t$, t and $t + \Delta t$:

$$q(\boldsymbol{x}, t) \leftarrow q(\boldsymbol{x}, t) + A[q(\boldsymbol{x}, t - \Delta t) - 2q(\boldsymbol{x}, t) + q(\boldsymbol{x}, t + \Delta t)]. \tag{2.91}$$

Unfortunately, in addition to damping out the computational mode introduced by the time level decoupling, this filter introduces unwanted numerical dispersion by slowing down the faster waves. Although this may be desirable in global climate models, turbulence simulations have been shown to be affected by the value chosen for the filter coefficient A. In particular, Polvani et al. [1994a] point out that the amplitude of the unbalanced motion is highly sensitive to changes in A. Surprisingly, little thought is usually given to the exact choice of this parameter. Here we present a thorough investigation of the effect of the filter coefficient in a simple 1D (i.e. no dependence on longitude) shallow water model using CASL.

The simulation is initialised with a zonal jet with a maximum velocity of $80ms^{-1}$ and a $120m$ depth anomaly (typical parameters for a mid-latitude jet). As the flow evolves, two waves propagate away from the jet, one to the north and one to the south. These waves are reflected back into the domain and interact with each other, producing complex wave patterns best illustrated by the divergence field. The time evolution of the depth and divergence fields is shown in figure 2.4 for three different values of A. The results in the top row, computed using $A = 0$,

i.e. no Robert-Asselin filtering, demonstrate the instability associated with the time level decoupling. Rapid oscillations develop in the divergence field and these grow and swamp the solution. The results on the bottom row were computed using a large (but typical) filter coefficient, $A = 0.1$. We can see that the effect of this is to almost completely wipe out the waves by time $t = 10$. While not having the same disastrous stability implications as the $A = 0$ case, this scenario is also undesirable. Not only does the filter cause numerical noise to be damped, it also damps the physical waves that form a valid part of the solution. We require the filter coefficient A to be large enough to prevent the oscillation due to the time level decoupling but small enough that it doesn't significantly affect the rest of the solution. In order to find this optimal value of A we define a parameter η that measures the change in Fr over the final 100 timesteps of the simulation:

$$\eta = \sqrt{\Sigma_{j=n-99}^{n} \left(\frac{\mathrm{Fr}_{j+1} - \mathrm{Fr}_j}{t_{j+1} - t_j} \right)^2}. \tag{2.92}$$

Rapid oscillations relating to time level decoupling will produce a large value of η, indicating that A needs to be increased. However, η will always be greater than zero as the Froude number is always changing so we have to choose a value that represents an acceptable amount of 'noise'. A few preliminary experiments showed that $\eta = 0.1$ was a sensible choice. Values larger than this prevented the algorithm from converging. Having specified this, we can now run a bisection search for the optimal value of A that allows this amount of noise. We compute this optimal value of A for a range of spatial resolutions n_ϕ and timesteps $\Delta t = \alpha \Delta t_{CFL}$ (where $\Delta t_{CFL} = \Delta\phi / c$ is the CFL timestep calculated using the gravity wave speed c and grid spacing $\Delta\phi$, and α is a constant in the range $0.1 \leq \alpha \leq 1.25$). This optimal value of A was used to compute the results in the middle row of figure 2.4.

Once we have found the optimal value $A = A_{opt}$ for each timestep Δt and

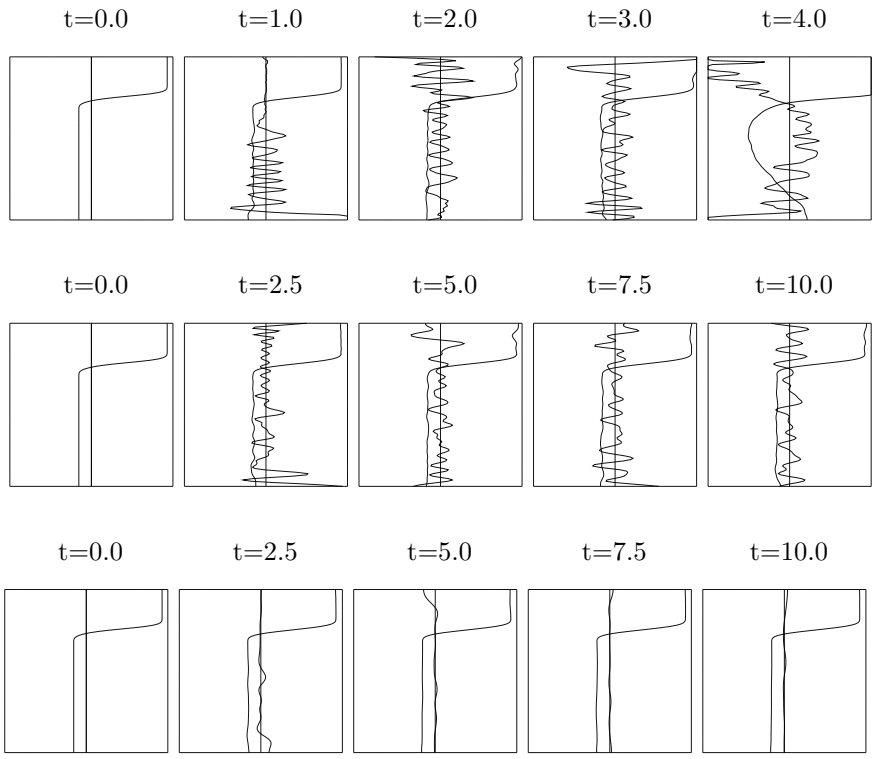


Figure 2.4: Evolution of the depth and divergence fields for $n_\phi = 128$ and $\Delta t = 0.5\Delta t_{CFL}$. The first row shows the results for $A = 0$, the middle row for $A = A_{opt} = 0.00424$ and the bottom row for $A = 0.1$. The maximum plotted value of $\tilde{h}$ is 0.1 and the maximum plotted value of δ is 0.5. Note the times given above each figure.

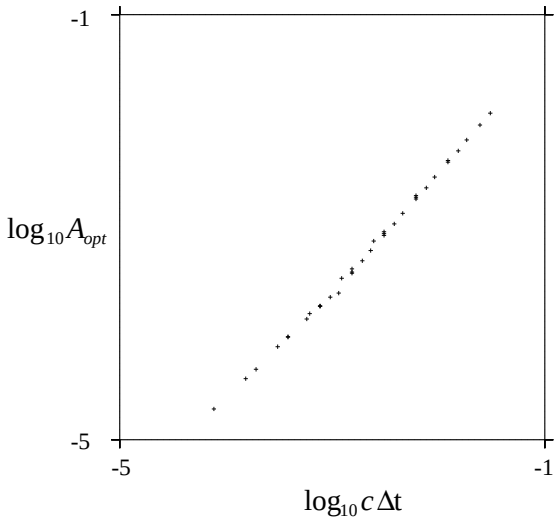


Figure 2.5: Log-log plot of the optimal filter coefficient A_{opt} versus $c\Delta t$ for all of the 1D simulations. Results from each value of n_ϕ, the spatial resolution, are plotted using the same symbol.

spatial resolution n_ϕ we can search for a relationship between this optimal value and the input parameters. The result, shown in figure 2.5, is that the filter coefficient is proportional to the gravity wave speed multiplied by the timestep and divided by the radius of the Earth (R_{Earth}):

$$A_{opt} \propto \frac{c\Delta t}{R_{Earth}}. \tag{2.93}$$

Faster wave speeds and larger timesteps require a larger filter coefficient to damp out the resulting noise. For the full spherical shallow water equations used in the rest of this thesis, we choose to set $A = 2c\Delta t$ ($R_{Earth} \equiv 1$) where the 'safety' factor of 2 was found to be sufficient to take into account the increased complexity of the system.

2.7.2 Tridiagonal, pentadiagonal and spectral convergence

We show here that for a discontinuity in PV, the formal accuracy of the spectral method is no greater than that of a 4th or even a 2nd order finite difference method.

We solve a simple second order differential equation

$$\psi_{xx} = q(x), \tag{2.94}$$

on the interval $-\pi \leq x \leq \pi$ with the (symmetry imposed) boundary condition that $\psi_x(-\pi) = \psi_x(\pi) = 0$ using three different methods: spectral, tridiagonal finite difference and pentadiagonal finite difference. The idea is to examine the order of accuracy of the three methods in the case where $q(x)$ is discontinuous. We take $q(x)$ to be a symmetric 'top-hat' profile with approximately unit width (adjusted so that the jumps occur midway between grid points). We remove the mean value of $q(x)$ so that the integral over the domain is zero. This gives the equation

$$\psi_{xx} = \begin{cases} -\bar{q} & -\pi < x < -x' \\ 1 - \bar{q} & -x' < x < x' \\ -\bar{q} & x' < x < \pi \end{cases}$$

where $\bar{q}$ is the mean value of q over the domain and x' is chosen to lie midway between grid points and such that $2x' \sim 1$. This equation can be solved analytically to give

$$\psi = \begin{cases} -\frac{\bar{q}}{2}x^2 - \bar{q}\pi x + C & -\pi < x < -x' \\ \frac{1}{2}(1-\bar{q})x^2 + C' & -x' < x < x' \\ -\frac{\bar{q}}{2}x^2 + \bar{q}\pi x + C & x' < x < \pi \end{cases} \tag{2.95}$$

where

$$C = -\frac{x'^3}{6\pi} - \frac{\pi x'}{3}, \tag{2.96}$$

and

$$C' = C + \frac{x'^2}{2}. \tag{2.97}$$

The constants in equation 2.95 are set by

a) imposing the zero derivative boundary conditions $\psi_x(-\pi) = \psi_x(\pi) = 0$ since the solution is even,

b) matching both ψ_x and ψ at $x = -x'$ and $x = x'$,

c) ensuring that $\bar{\psi} = 0$.

Spectral approach

The spectral method requires us to transform equation 2.94 into Fourier space. We write ψ and q in terms of their Fourier series and substitute into 2.94:

$$-\Sigma a_k k^2 e^{ikx} = \Sigma b_k e^{ikx}, \tag{2.98}$$

where a_k and b_k are the Fourier coefficients of ψ and q respectively. Equating the Fourier coefficients term by term gives us

$$a_k = -\frac{b_k}{k^2}. \tag{2.99}$$

We now compute $\psi(x)$ by transforming back into physical space. Repeating this algorithm for $k = 2^n$ with n taking each value between 4 and 11 will give us a range of solutions to compare with the exact solution and we can then calculate how the accuracy of the procedure depends on n.

Tridiagonal approach

Discretising 2.94 using centred finite differences gives

$$\frac{\psi_{j+1} - 2\psi_j + \psi_{j-1}}{(\Delta x)^2} = q_j, \tag{2.100}$$

for $j = 2, n-1$ where j is the index of the grid point and Δx is the grid spacing. For the boundary points we specify homogeneous boundary conditions $\psi(1) = \psi(n) = 0$. ψ is afterwards adjusted to have zero mean. This method can be shown to be second order accurate when $\psi(x)$ is twice continuously differentiable. Equation 2.100 gives us a matrix equation

$$\begin{pmatrix} b_1 & c_2 & & & & \\ a_2 & b_2 & c_3 & & & \\ & a_3 & b_3 & c_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} & c_n \\ & & & & a_n & b_n \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_{n-1} \\ \psi_n \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_{n-1} \\ q_n \end{pmatrix} \tag{2.101}$$

where

$$a_i = c_i = \frac{1}{\Delta x^2}, \tag{2.102}$$

$$b_i = -\frac{2}{\Delta x^2}. \tag{2.103}$$

The matrix equation is easily inverted in two steps using the Thomas algorithm [Press et al., 1992]. First we sweep forwards in i to calculate

$$\hat{\psi}_2 = \frac{q_2}{d_2} \tag{2.104}$$

$$\hat{\psi}_i = \frac{q_i - a_i\psi_{i-1}}{d_i}, \qquad i = 3, \ldots, n \tag{2.105}$$

where

$$d_2 = b_2 \tag{2.106}$$

$$d_i = b_i - \frac{a_i c_{i-1}}{d_{i-1}}, \qquad i = 3, \ldots, n \tag{2.107}$$

This is equivalent to performing an LU decomposition and forward substitution. The second step is to backsubstitute to obtain

$$\psi_i = \hat{\psi}_i - \frac{a_{i+1}}{d_i}\hat{\psi}_{i+1}, \qquad i = n-1, \ldots, 2. \tag{2.108}$$

Pentadiagonal approach

Lele [1992] describes a range of finite differencing schemes, including the following fourth order, pentadiagonal scheme.

$$\begin{pmatrix} c_1 & d_1 & e_1 & & & \\ b_2 & c_2 & d_2 & e_2 & & \\ a_3 & b_3 & c_3 & d_3 & e_3 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & a_{n-1} & b_{n-1} & c_{n-1} & d_n \\ & & & a_n & b_n & c_n \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_{n-1} \\ \psi_n \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_{n-1} \\ q_n \end{pmatrix}, \tag{2.109}$$

where

$$\left.\begin{array}{rcl} a_i = e_i &=& \frac{3}{44\Delta x^2} \\ b_i = d_i &=& \frac{12}{11\Delta x^2} \\ c_i &=& \frac{-51}{22\Delta x^2} \end{array}\right\} i = 2, \ldots n-2 \tag{2.110}$$

and

$$\begin{array}{rcl} a_1 = b_1 &=& 0 \\ c_1 &=& -27\Delta x^2 \\ d_1 &=& 15\Delta x^2 \\ e_1 &=& -\Delta x^2 \\ a_{n-1} &=& -\Delta x^2 \\ b_{n-1} &=& 15\Delta x^2 \\ c_{n-1} &=& -27\Delta x^2 \\ d_{n-1} = e_{n-1} &=& \frac{3}{44\Delta x^2}. \end{array} \tag{2.111}$$

Equation 2.111 enforces the boundary conditions.

This pentadiagonal matrix equation is solved in a similar method to the tridiagonal version. The method consists of two forward sweeps to eliminate the a_i and b_i and back substitution to calculate the solution ψ_i.

Results

To compute the order of accuracy of the above methods we run them using different resolutions $n = 2^p, p = 4, \ldots, 11$ where n is the number of grid points. For each run we calculate the L_1 and L_2 norms of the error:

$$L_1 = \frac{1}{n}\Sigma_{i=1}^{n}|\psi_{ex}(x_i) - \psi_i^{\text{num}}|, \tag{2.112}$$

$$L_2 = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(\psi_{ex}(x_i) - \psi_i^{\text{num}})^2}\,, \tag{2.113}$$

where $\psi_{ex}(x_i)$ is the exact solution at $x = x_i$ and ψ_i^{num} is the numerical solution computed by a particular method. The results are plotted in figure 2.6. We see that the order (given by the slope of the plotted line) of all three schemes is the same: second. Although using the pentadiagonal instead of the tridiagonal scheme reduces the magnitude of the error (the ratio of the errors is approximately 3-3.5), the difference is not substantial.

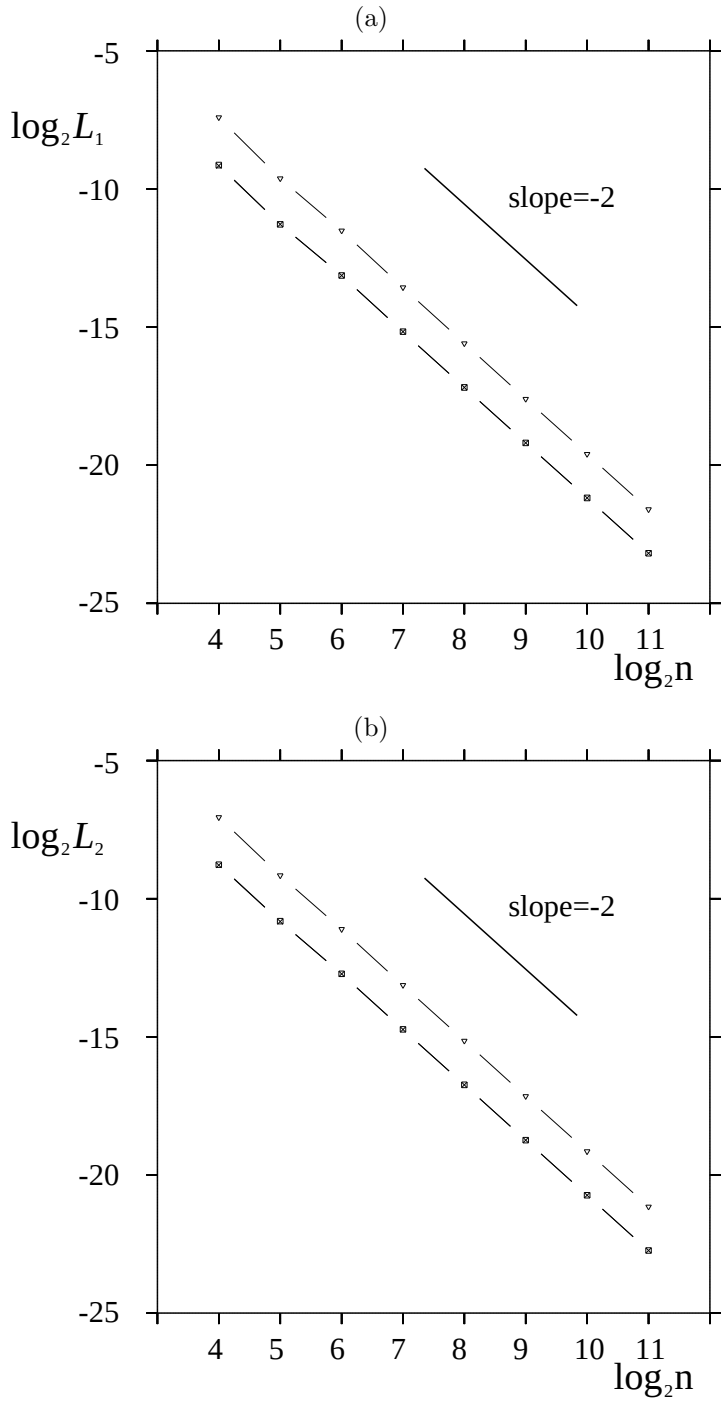


Figure 2.6: The L_1 (a) and L_2 (b) error measures of the spectral (cross-dash-cross line), tridiagonal (triangle-dash-triangle line) and pentadiagonal (square-dash-square line) plotted against the number of gridpoints n_ϕ. All axes are log scaled.

Chapter 3

Spherical shallow water turbulence

3.1 Introduction

In this chapter we present the results of around 150 simulations of shallow water turbulence. The properties of the flow are controlled by their Froude and Rossby numbers, $\mathrm{Fr} \equiv |\boldsymbol{u}|/c(1+\tilde{h})^{\frac{1}{2}}$ and $\mathrm{Ro} \equiv \zeta/2\Omega$. The simulations described in this chapter span the Fr-Ro parameter space (see figure 3.3). First we describe the initial conditions and recap the model settings. We then focus on four cases from the extreme corners of our parameter space. We examine the evolution of the flow and observe how various properties such as cyclone-anticyclone asymmetry and jet formation vary between simulations.

3.1.1 Initial conditions

The initial conditions are generated by adding a random, isotropic perturbation to a zonal PV distribution (see figure 3.1). The perturbation is constructed

by generating random numbers at the gridpoints which are then correlated over a length scale L_C which determines the initial size of coherent structures. The strength of the turbulence, or of the circulation, is controlled by the r.m.s. PV anomaly ϖ_{rms}, where $\varpi = \Pi - f$. Figure 3.2 shows the effect of changing these two parameters on the configuration of the perturbation to the background zonal flow. The (polar) Rossby radius $L_D = c/2\Omega$ (with $c = \sqrt{gH}$) is also specified. This setup enables us to ensure that the important length scales of flow are resolved. We take $\Omega = 2\pi$ so that a unit of time corresponds to one day. The basic grid resolution is $n_\lambda = n_\phi = 128$, and the time step is chosen to lie between half and one times the CFL time step, $\Delta\phi/c$, where $\Delta\phi = \pi/n_\phi$ is the latitude grid spacing, for high temporal accuracy (Smith & Dritschel 2006). The PV anomaly is ramped up over a 5-day ramp period ($\Delta\tau_I = 5$), during which significant advection of the original PV contours occurs. The ramp function is given by

$$T(\tau) \equiv \tfrac{1}{2}[1 - \cos(\pi\tau/\Delta\tau_I)]\,, \tag{3.1}$$

where $0 \le \tau \le \Delta\tau_I$ is a fictitious ramping time. Note: $T(0) = 0$, $T(\Delta\tau_I) = 1$, $T'(0) = T'(\Delta\tau_I) = 0$. This establishes the initial flow fields at '$t = 0$' in a state of near balance [Viúdez and Dritschel, 2004a].

The circulation regime is perhaps best characterised by the Froude and (polar) Rossby numbers, $\mathrm{Fr} \equiv |\boldsymbol{u}|/c(1+\tilde{h})^{\frac{1}{2}}$ and $\mathrm{Ro} \equiv \zeta/2\Omega$. These cannot be easily predicted from the specified input parameters and are also to some extent dependent on the initial spatial configuration of PV contours, especially at higher Froude numbers. However, having observed that $L_C/L_D \sim \mathrm{Fr}/\mathrm{Ro}$ and that increasing ϖ increases both the Rossby and Froude number, is it possible to estimate where a run will lie, at least initially, in the Fr–Ro parameter space. During the nonlinear flow evolution, the Froude and Rossby numbers may change significantly from their initial values so we use instead the time-mean (over $t = 0$–40) rms average over the domain $(\mathrm{Fr}_{rms}, \mathrm{Ro}_{rms})$. We have performed 148 simulations throughout

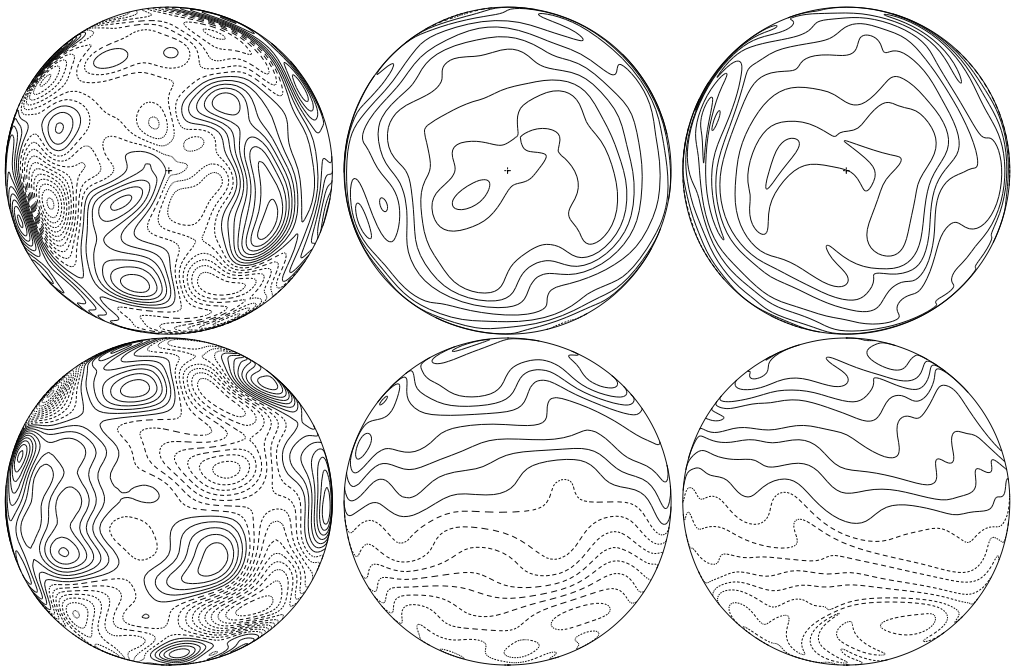


Figure 3.1: An example of the initial PV contours. The top row shows the polar view and the bottom row shows the equatorial view. The first column is an example of the random isotropic perturbation which is added onto a pure zonal flow. The sum of the perturbation and zonal flow is shown in the second column. The third column shows the PV contours after the ramped initialisation.

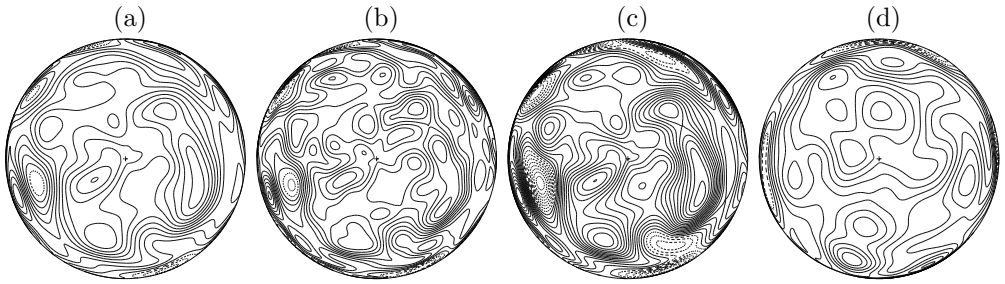


Figure 3.2: A polar view of the random perturbation computed for different parameters L_C, ϖ_{rms} and random number seed. Comparing figure (b) with figure (a) shows that halving L_C results in more small scale features; comparing figure (c) with figure (a) shows that multiplying ϖ by 4 steepens the PV gradients; and comparing figure (d) with figure (a) shows that using a different random number seed gives an entirely different spatial configuration of the PV contours. The contour interval is the same in all four figures.

the parameter space (see figure 3.3), focusing on the regime with high Froude number as it is here where the numerical method and the concept of balance will be tested most severely.

Four simulations are described next, one from each corner of the parameter space (see figure 3.3). In cases 2 and 4 the turbulence peaks in complexity (measured by the number of nodes representing the PV contours) at $t = 1$ while in case 3 the turbulence peaks in complexity around $t = 6$, whereas in case 1 the complexity increases slowly throughout the 40 day run. For ease of comparison, each simulation starts with the same spatial structure of ϖ. All other parameters are given in Table 3.1. All simulations have been run for 40 days.

3.1.2 Flow evolution

The shallow water equations exactly conserve mass, energy and angular momentum. By construction, the CASL model conserves mass. In figure 3.4 we plot

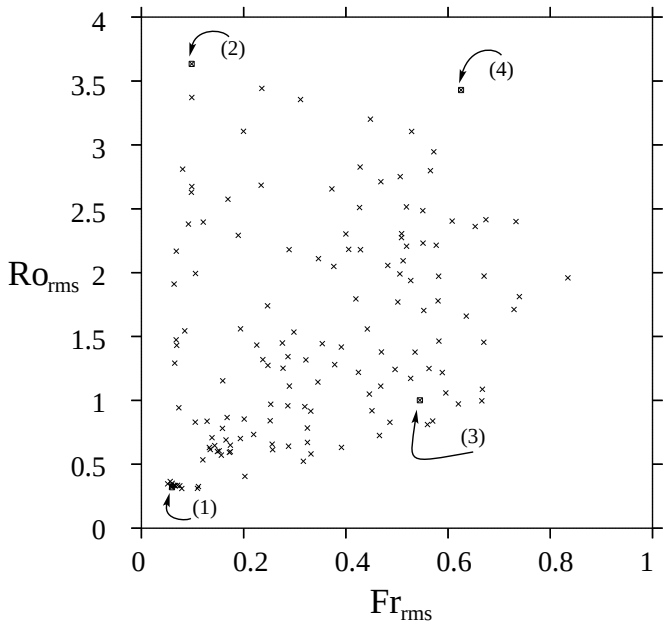


Figure 3.3: Crosses mark the location of the simulations in the Fr–Ro parameter space. Squares mark the four simulations focused on in the following section.

	L_D	L_C	ϖ	$\mathrm{Fr}_{\mathrm{max}}$	$\mathrm{Ro}_{\mathrm{max}}$	$\mathrm{Ro}_{\mathrm{min}}$
(1)	0.883	0.147	0.100	0.06	0.32	-0.33
(2)	0.400	0.100	1.000	0.10	3.63	-2.04
(3)	0.294	0.100	0.490	0.54	1.00	-0.94
(4)	0.883	0.147	1.250	0.63	3.43	-1.88

Table 3.1: The first three columns give the input parameters for each case. The last three columns give the time averaged Froude and Rossby numbers over the 40 day run.

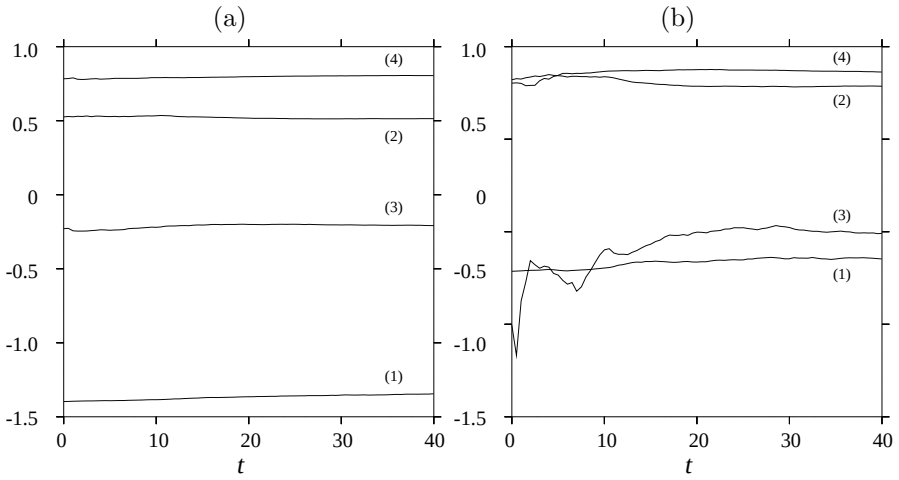


Figure 3.4: Time evolution of the log (base 10) of (a) total energy E and (b) angular momentum J in the four simulations.

the time evolution of the other two quantities (per unit area)

$$E = \frac{1}{8\pi}\iint \left[(1+\tilde{h})|\boldsymbol{u}|^2 + c^2\tilde{h}^2\right] \mathrm{d}S \tag{3.2}$$

$$J = \frac{1}{4\pi}\iint \left[(1+\tilde{h})r(u+\Omega r) - \Omega r^2\right] \mathrm{d}S \tag{3.3}$$

where $r = \cos\phi$ and $\mathrm{d}S = r\mathrm{d}\phi\mathrm{d}\lambda$. Note that E and J, defined in this way, vanish for $\tilde{h} = 0$ and $\boldsymbol{u} = \mathbf{0}$. These definitions do not include the constants that typically account for a significant majority (90-99%) of the total [Mohebalhojeh and Dritschel, 2007]. The energy is well conserved, varying by only 5-10 percent. The angular momentum varies more, up to 18 percent in case 4, but since, unlike the energy, this quantity is not sign definite, its variation is less significant. These quantities are better conserved in higher resolution CASL models [Mohebalhojeh and Dritschel, 2007], but then it is impractical to perform the wave–vortex decomposition described in the next section.

Figure 3.5 shows the time evolution of $\mathrm{Fr}_{\max}$, $\mathrm{Ro}_{\min}$ and $\mathrm{Ro}_{\max}$ (sampled every time step) for the four simulations. The maximum Froude number varies only

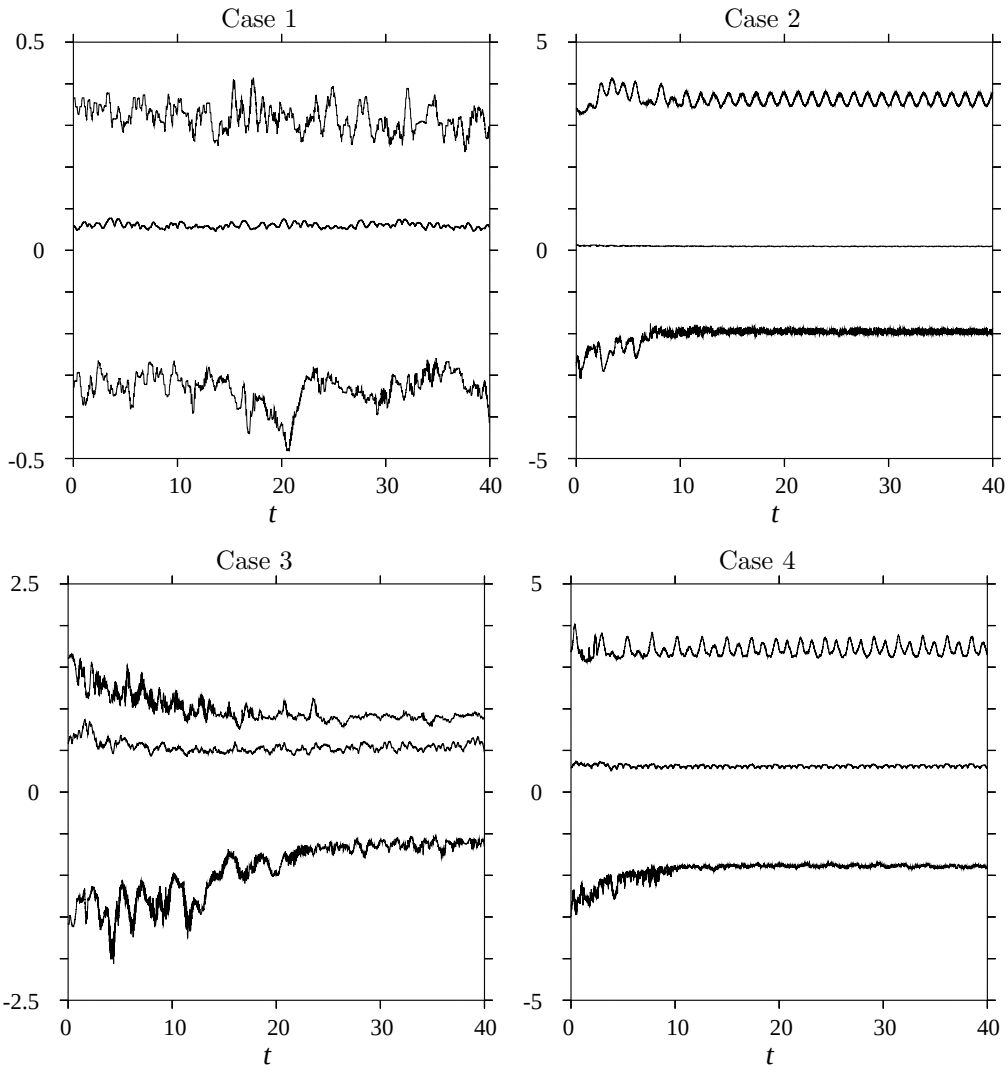


Figure 3.5: Time evolution of maximum Froude number (middle, bold solid curve) and minimum and maximum Rossby numbers (thin solid curves).

weakly while the maximum and minimum Rossby numbers display a combination of low- and high-frequency oscillations. The high frequency, sub-inertial oscillations can be seen particularly clearly in cases 2 and 4. The time averaged values of $\mathrm{Fr}_{\max}$, $\mathrm{Ro}_{\min}$ and $\mathrm{Ro}_{\max}$ are given in table 3.1.

The real complexity of these flows is revealed by the PV field, shown at times $t = 0$, 5 and 10 for the four cases in figures 3.6 – 3.9. Note that for cases 2, 3 and 4 the contours are so dense that not all of them are plotted (see figure captions for details). Also shown are the corresponding depth $\tilde{h}$, divergence δ and ageostrophic vorticity γ fields. Each rectangle corresponds to the full domain, $-\pi \leq \lambda \leq \pi$ and $-\frac{1}{2}\pi \leq \phi \leq \frac{1}{2}\pi$. The flows range from wavy-turbulence, consisting of nonlinear waves and few vortices except near the poles, to strong turbulence, with many active vortices entangling the background PV field associated with the planetary rotation. This complexity is the result of repeated vortex interactions and breaking nonlinear Rossby waves, and is typical of weakly-dissipative flows.

The tendency in all these flows is for vortices of the same sign to merge and migrate polewards or equatorwards depending on their sign. The properties of these vortices will be explored in section 3.2. A striking example of vortex merger may be seen in the final two frames of figure 3.7. In the PV field, two cyclonic vortices combine and spin off an encircling filament. In the height field, this shows up as the merger of two depressions, while in the γ field, the vortices appear as the strongest ageostrophic features over the entire globe. Even in the δ field, a weak quadrupolar trace is evident around each vortex. Features in each field can be traced back to features in the PV field. This indicates that all fields, even δ and γ, are dominated by the balanced, PV induced, component of the flow. We shall demonstrate this more rigorously in section 4.2.

All simulations also exhibit an equatorial wave (however of global scale), particularly in cases 2 and 4 (see figures 3.7 and 3.9). This wave induces significant

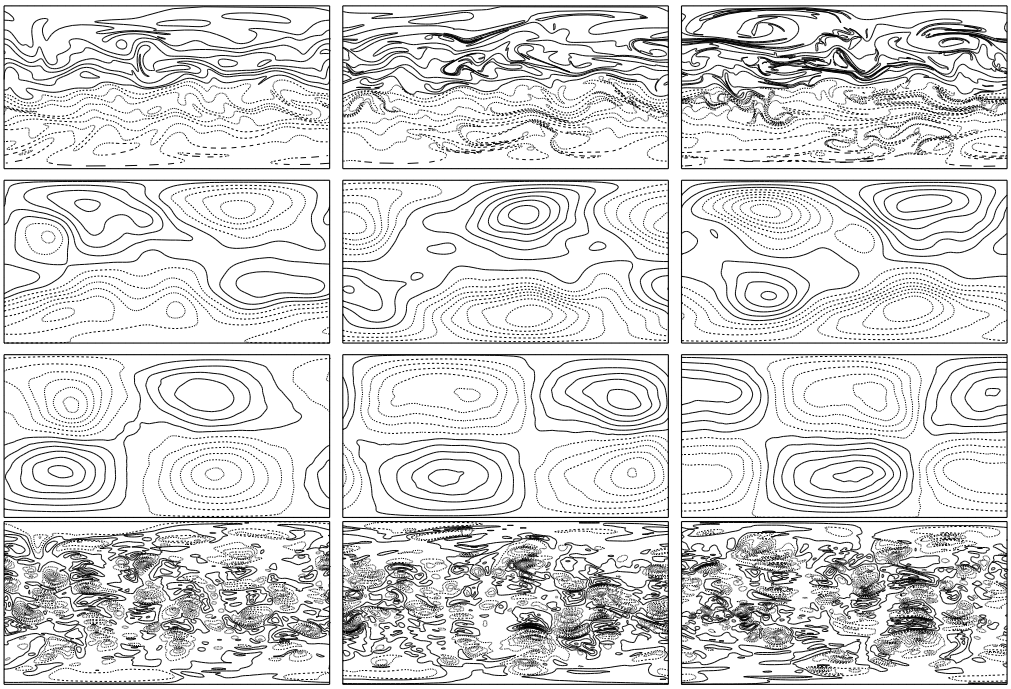


Figure 3.6: Evolution of Π (top row), $\tilde{h}$ (second row), $\delta/(2\Omega)$ (third row) and $\gamma/(2\Omega)^2$ (bottom row) fields for case 1, shown at times $t = 0$, 5, 10. The contour interval is $\pi/2$ for Π, 0.002 for $\tilde{h}$, 0.0005 for $\delta/(2\Omega)$ and 0.005 for $\gamma/(2\Omega)^2$.

cross equatorial flow and north-south excursions of the PV contours (most clearly seen in movies of the flow evolution). This results in large displacements of the 'polar' vortex or vortices. We shall see in the following section the effect that this has on the dynamics of any jets that form.

3.2 Cyclone-anticyclone asymmetry

Cyclone-anticyclone asymmetry has long been observed in atmospheric data [Venn, 1887], laboratory experiments [Perret et al., 2006] and numerical simulations [Polvani et al., 1994a, Theiss, 2004, Stegner and Dritschel, 2000, and references therein]. Polvani et al. [1994a] found a skewness (as defined in equa-

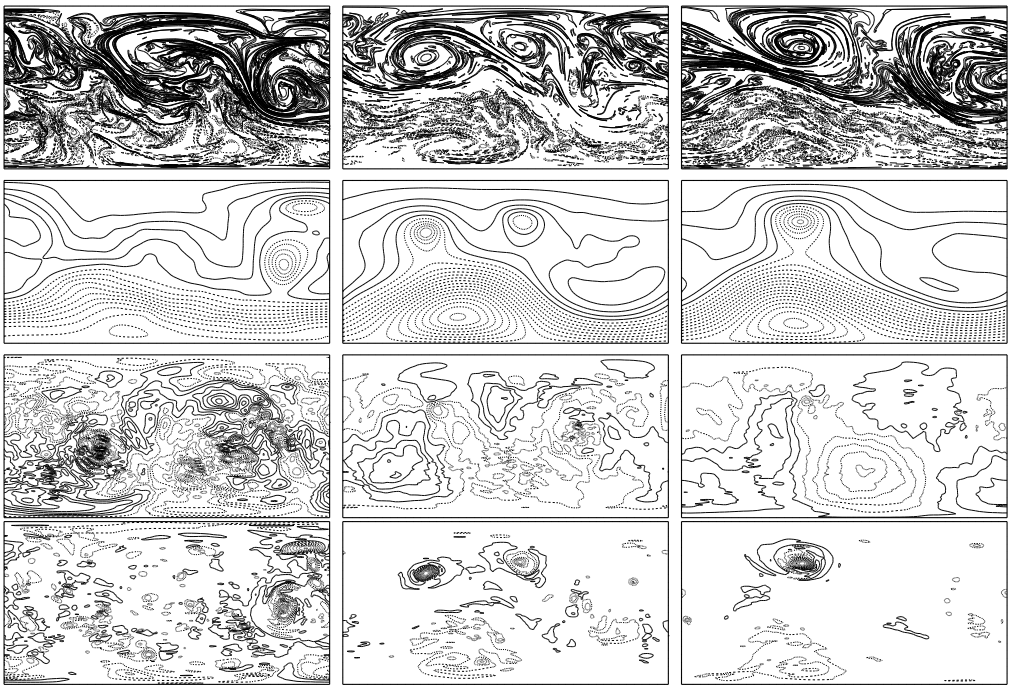


Figure 3.7: Evolution of Π (top row), $\tilde{h}$ (second row), $\delta/(2\Omega)$ (third row) and $\gamma/(2\Omega)^2$ (bottom row) fields for case 2, shown at times $t = 0$, 5, 10. Only every $8th$ contour of the PV field is plotted. The contour interval is 4π for Π, 0.001 for $\tilde{h}$, 0.0005 for $\delta/(2\Omega)$ and 0.5 for $\gamma/(2\Omega)^2$.

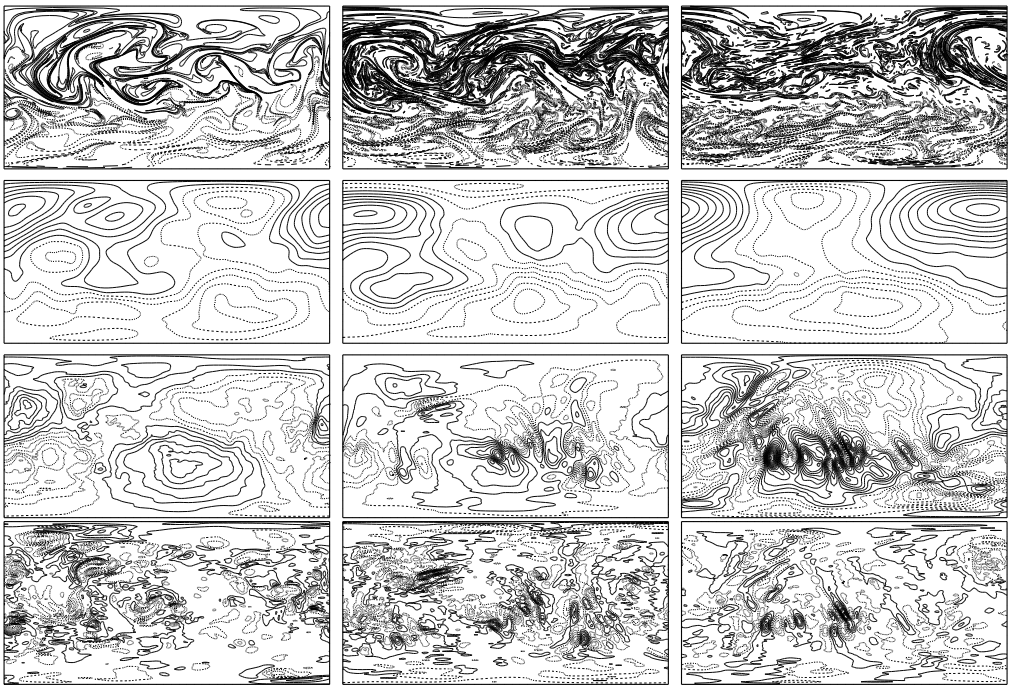


Figure 3.8: Evolution of Π (top row), $\tilde{h}$ (second row), $\delta/(2\Omega)$ (third row) and $\gamma/(2\Omega)^2$ (bottom row) fields for case 3, shown at times $t = 0$, 5, 10. Only every $5th$ contour of the PV field is plotted. The contour interval is $5\pi/2$ for Π, 0.05 for $\tilde{h}$, 0.01 for $\delta/(2\Omega)$ and 0.1 for $\gamma/(2\Omega)^2$.

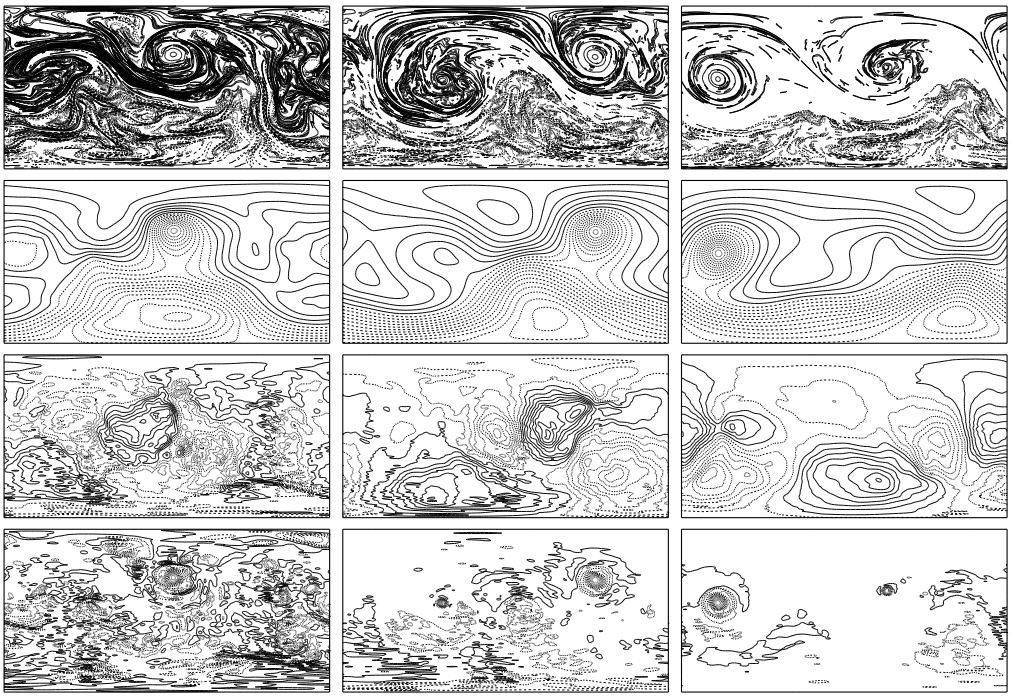


Figure 3.9: Evolution of Π (top row), $\tilde{h}$ (second row), $\delta/(2\Omega)$ (third row) and $\gamma/(2\Omega)^2$ (bottom row) fields for case 4, shown at times $t = 0$, 5, 10. Only every $8th$ contour of the PV field is plotted. The contour interval is 4π for Π, 0.02 for $\tilde{h}$, 0.01 for $\delta/(2\Omega)$ and 0.5 for $\gamma/(2\Omega)^2$.

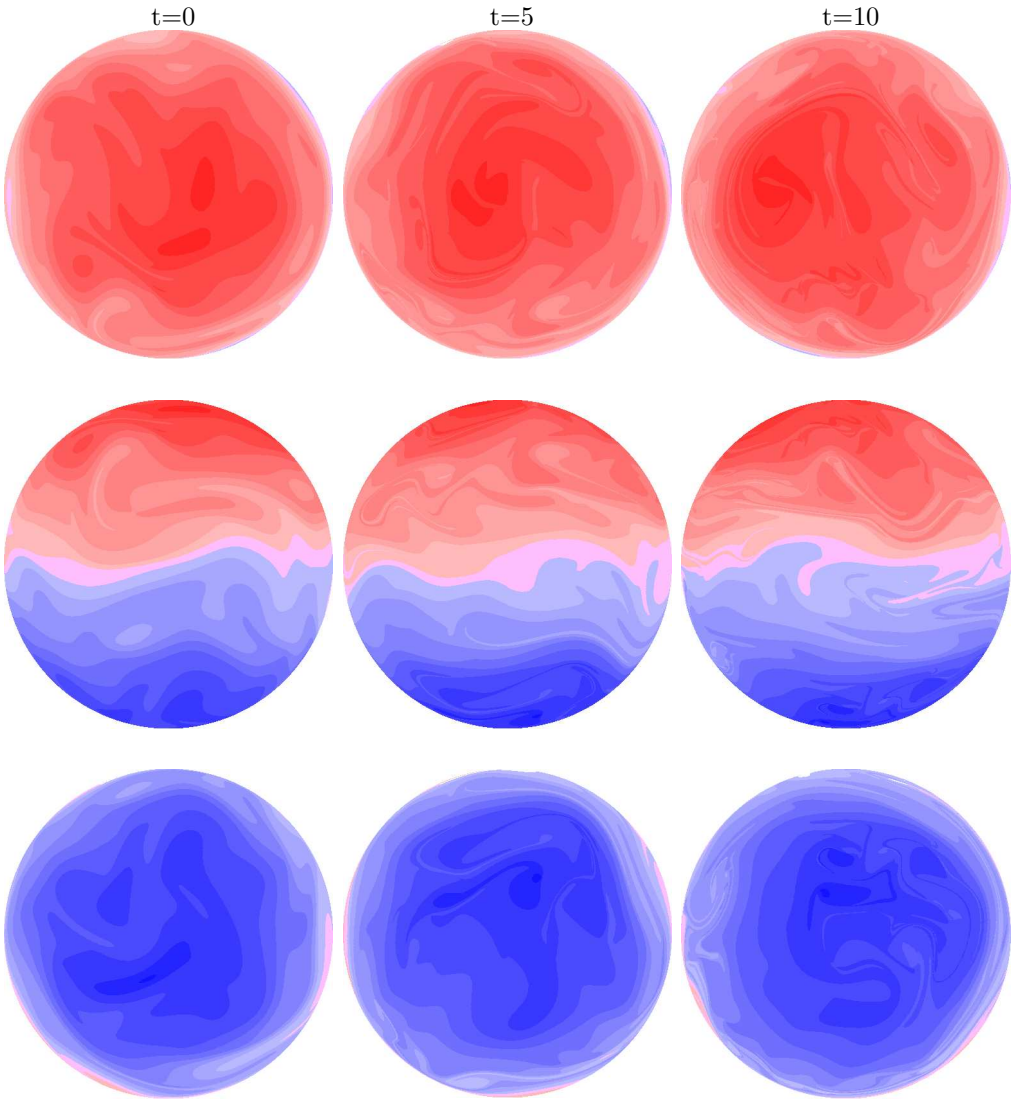


Figure 3.10: Evolution of Π for case 1, shown from the north pole (top row), from the equator (middle row) and from the south pole (bottom row). Red indicates positive Π, blue indicates negative Π.

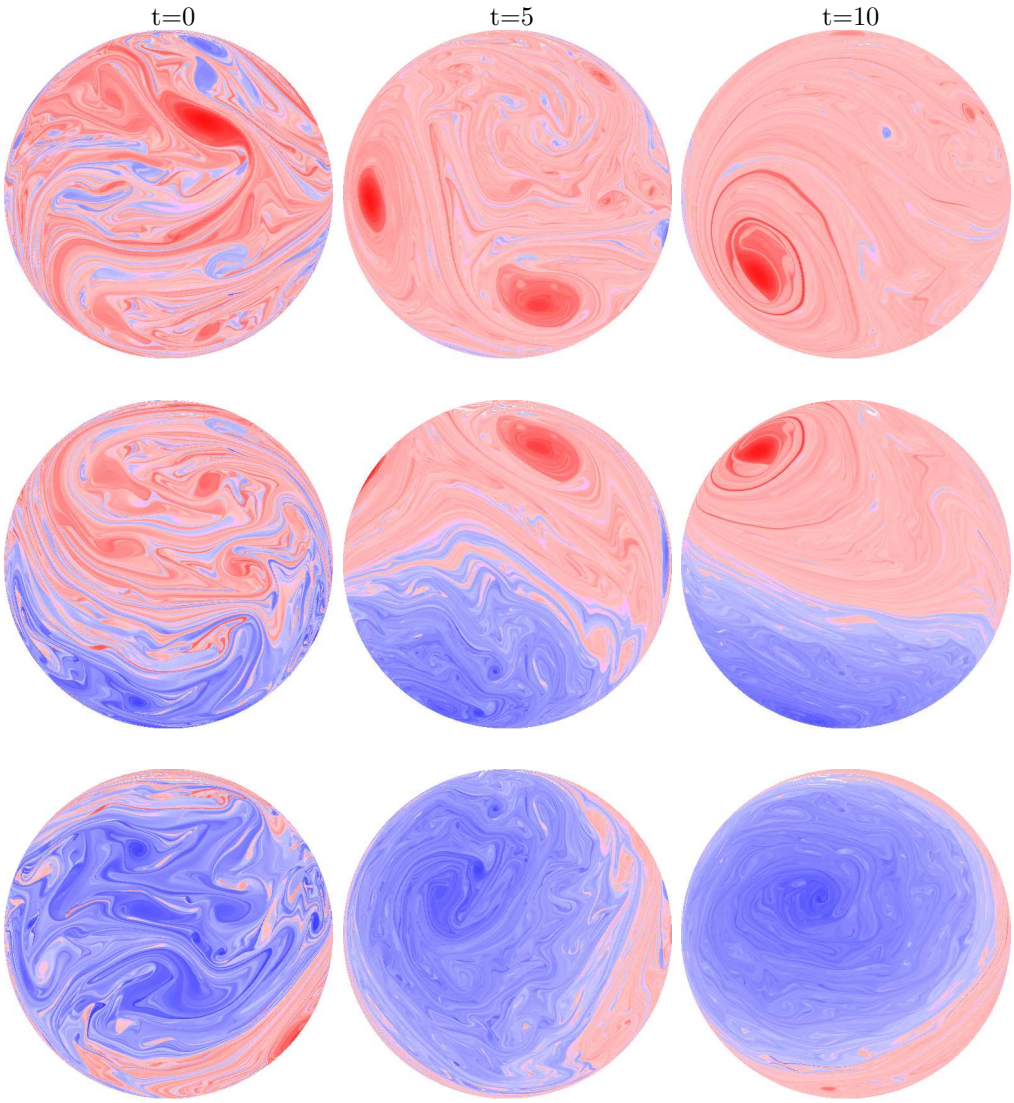


Figure 3.11: Evolution of Π for case 2, shown from the north pole (top row), from the equator (middle row) and from the south pole (bottom row). Red indicates positive Π, blue indicates negative Π.

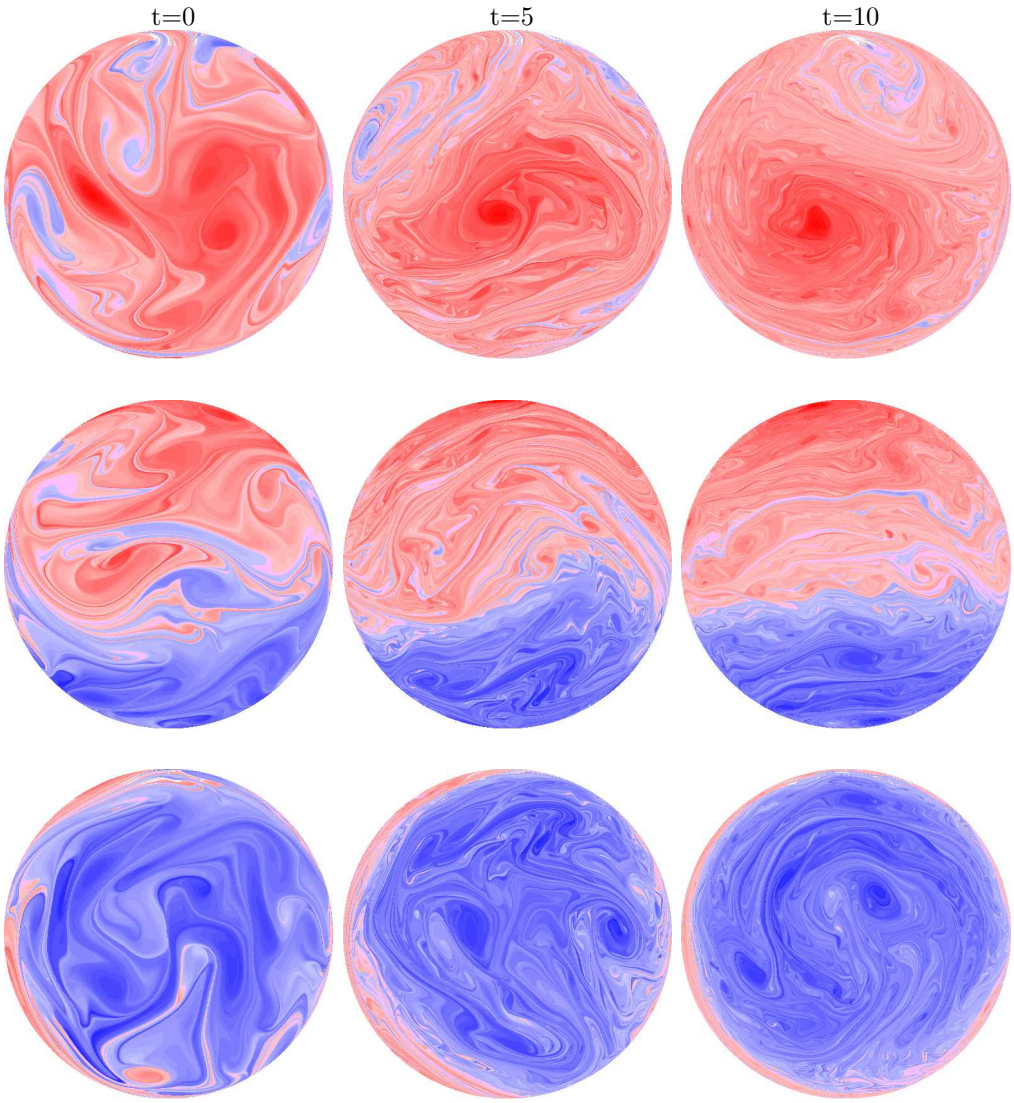


Figure 3.12: Evolution of Π for case 3, shown from the north pole (top row), from the equator (middle row) and from the south pole (bottom row). Red indicates positive Π, blue indicates negative Π.

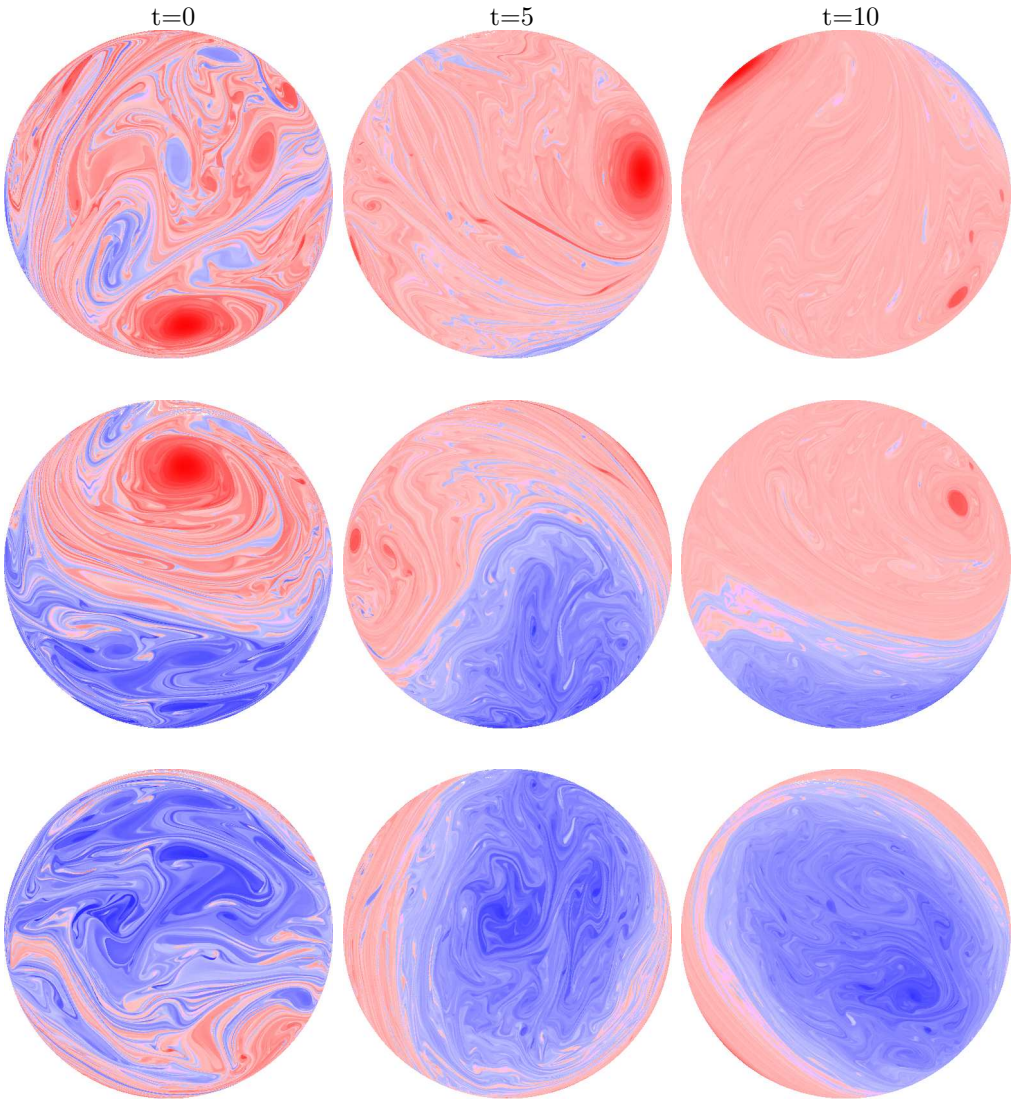


Figure 3.13: Evolution of Π for case 4, shown from the north pole (top row), from the equator (middle row) and from the south pole (bottom row). Red indicates positive Π, blue indicates negative Π.

tion 3.4) in favour of anticyclonic vorticity with the asymmetry becoming more pronounced with increasing Froude number. They also note the appearance of some large-scale cyclonic structures, thus demonstrating that, contrary to the suggestion of Cushman-Roisin and Tang [1990], cyclones do not just fall apart. In fact, as demonstrated below, cyclones, while fewer, tend to be much stronger than anticyclones.

Following Polvani et al. [1994a], we measure the cyclone-anticyclone asymmetry by computing the skewness

$$S(q) \equiv \frac{\langle q^3 \rangle}{\langle q^2 \rangle^{3/2}}, \tag{3.4}$$

where $\langle . \rangle$ denotes the spatial average, of various fields q. Polvani et al. [1994a] examined $S(\zeta)$, which is inappropriate here since $\zeta > 0$ does not correspond to cyclonic circulation in both hemispheres. Instead, we examine $S(\tilde{h})$ and $S(\gamma)$, both of which indicate cyclonic circulation for $S > 0$ and anticyclonic circulation for $S < 0$. The time mean values over $20 \leq t \leq 40$ of $S(\tilde{h})$ and $S(\gamma)$ are tabulated in table 3.2. The first feature to note is that case 3 is the only case for which the average skewness is positive for either field. Case 1, as expected from its location in the Fr-Ro parameter space, exhibits the behaviour expected of a near-geostrophically balanced flow, in that the skewness stays close to, albeit less than, zero for the length of the run (not shown). In cases 2 and 4 asymmetry appears early on and levels out after about 10 days.

Another way to visualise the asymmetry is to examine the probability density function (PDF) of these fields. The depth field is shown in figure 3.14 (γ is qualitatively similar). As first noted by Venn (1887), departures from the mean surface pressure in atmospheric data do not form a Gaussian distribution. Here, in shallow water flows, the depth field plays the role of surface pressure through hydrostatic balance. Again, case 1 exhibits the behaviour expected of a near-

	$S(\tilde{h})$	$S(\gamma)$
(1)	-0.19	-0.40
(2)	-0.40	-3.14
(3)	0.08	0.51
(4)	-0.26	-1.93

Table 3.2: Time averaged skewness of $\tilde{h}$ and γ over the last half of each simulation.

geostrophically balanced flow, in that its PDF is the most symmetric although even in this case there is a bias towards high pressure regions. The PDFs for cases 2 and 4 are significantly more skewed but both also exhibit a significant tail of low pressure values indicating that, although high pressure regions dominate, low pressure regions tend to be much more intense. This is the key observation made by Venn (1887) regarding surface pressure. This pattern is reversed in case 3. Looking at the final depth field for all four cases (figure 3.15), we see that, despite the skewness in favour of cyclonic circulation, the dominant feature in case 3 is a large, intense anticyclone which accounts for the PDF distribution in figure 3.14. This case of low Ro and high Fr may be the least representative of global atmospheric motion.

We now examine the asymmetry over the Fr-Ro parameter space. In figure 3.16 we plot the skewness $S(\tilde{h})$ against Fr and Ro. For the majority of the simulations the skewness is negative, indicating dominance of anticyclonic circulation. However, at low Rossby number and, to a lesser extent, high Froude number, some simulations have a strong positive skewness. As we saw above, the skewness does not reveal the whole picture and it is possible that the flows are in fact dominated by a few extreme vortices of the sign opposite to that which the skewness suggests. To investigate this we plot the ratio $|\tilde{h}_{max}/\tilde{h}_{min}|$ against Fr and Ro (see figure 3.17). This reveals that in the majority (57%) of simulations $|\tilde{h}_{max}| < |\tilde{h}_{min}|$, i.e. cyclones are more intense than anticyclones in these flows.

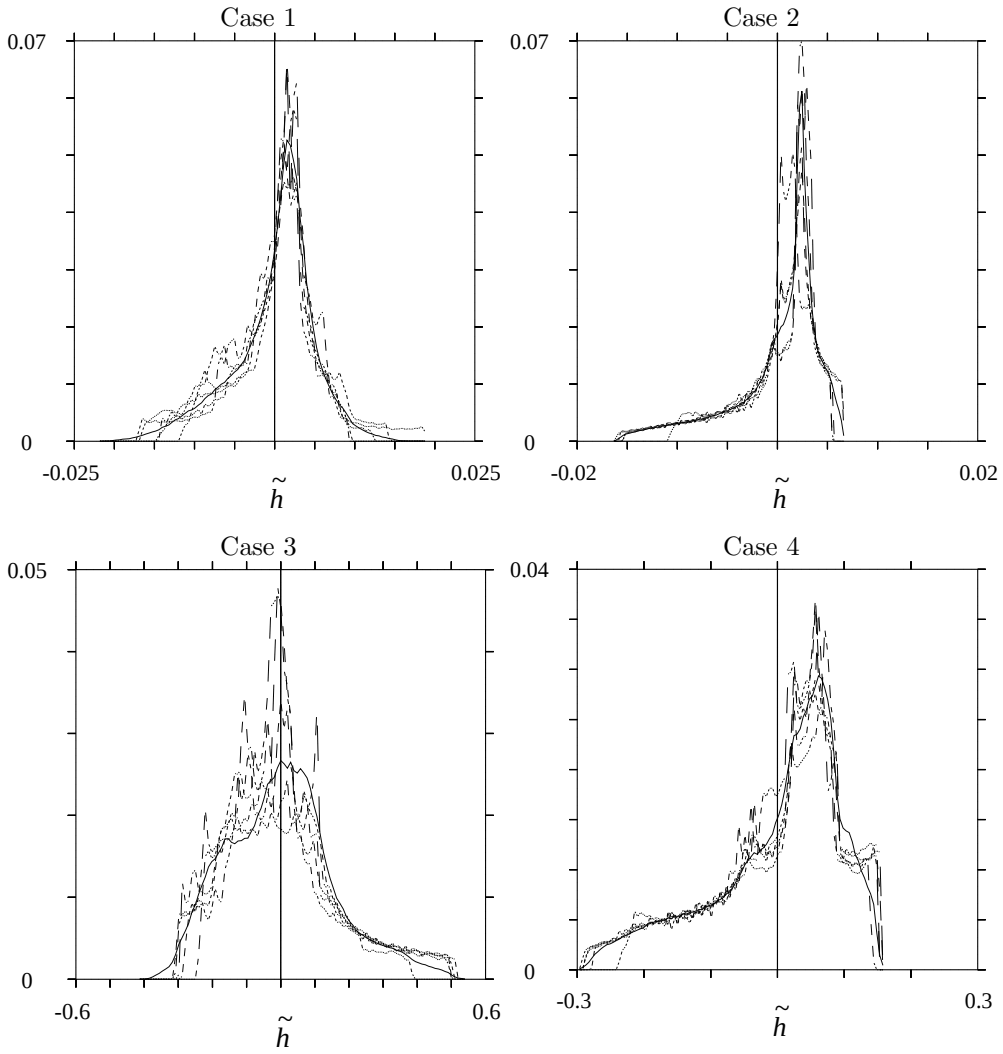


Figure 3.14: Probability density functions of depth $\tilde{h}$ for the 4 cases. Dotted lines show the distribution at $t = 0, 10, 20, 30, 40$ and the solid line shows the time-mean distribution.

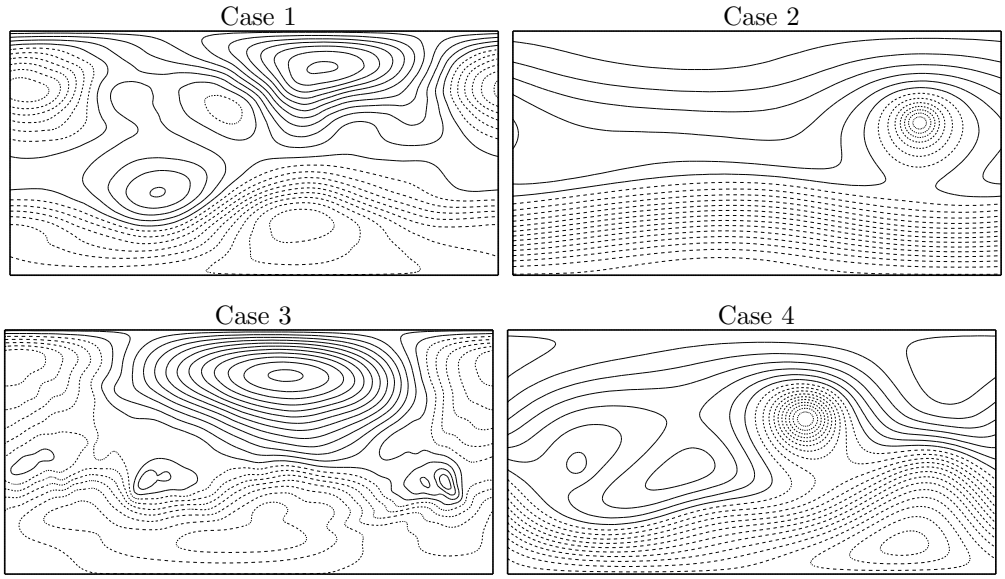


Figure 3.15: Depth field at $t = 40$ for all four cases, plotted with 20 contour intervals spanning 0 to $\tilde{h}_{max}$.

Although 57% is on the margin of being significant (based on a rough estimate of the likely error in the 148 cases), this suggests that Venn's observations of real data are consistent with shallow water turbulence over much of the Fr-Ro parameter space.

3.3 Potential vorticity homogenisation

Regions of homogenised potential vorticity are found in the oceans (e.g. ocean gyres [Rhines and Young, 1982]), in the Earth's atmosphere (e.g. the 'surf zone' surrounding the polar vortex [Polvani et al., 1994b]), and in the atmospheres of the gaseous outer planets (e.g. the bands between Jupiter's alternating zonal jets [Marcus and Lee, 1998]). The process of PV homogenisation is an important geophysical mechanism, responsible for the sharpening of jets through the expulsion of PV gradients to the edge of flow structures or domains [Rhines, 1994].

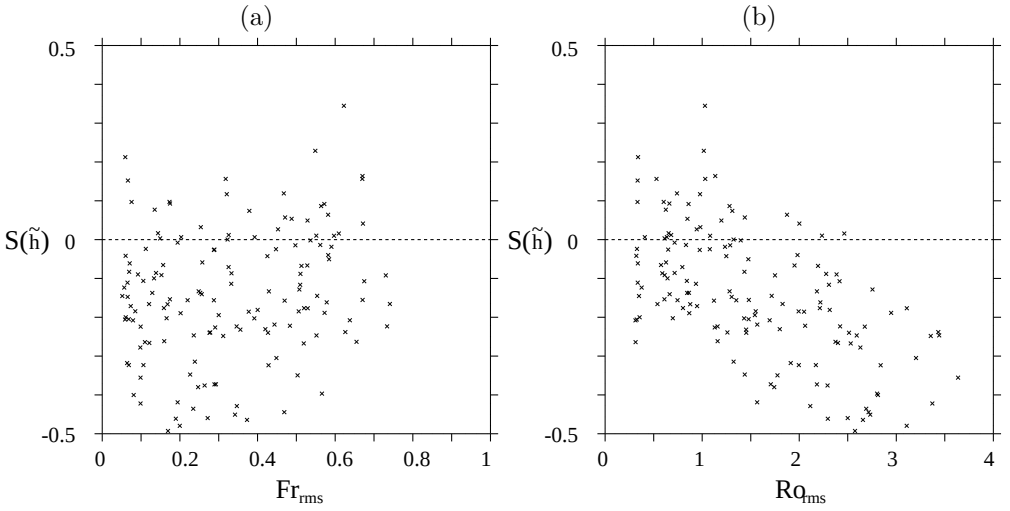


Figure 3.16: Skewness of $\tilde{h}$ for each simulation plotted against Fr(a) and Ro(b).

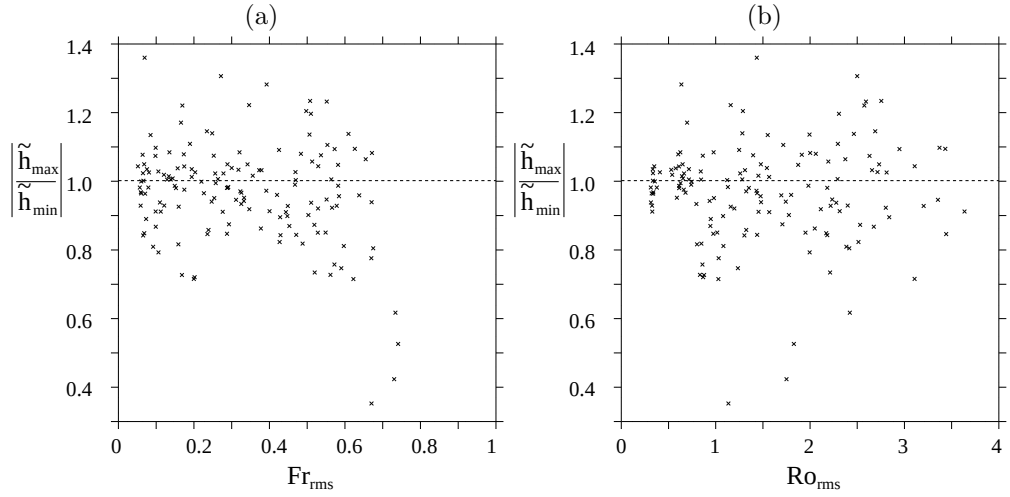


Figure 3.17: Absolute value of $\tilde{h}_{max}/\tilde{h}_{min}$ for each simulation plotted against Fr(a) and Ro(b).

A recent study [Dritschel and McIntyre, 2008] presents a comprehensive review of the lab experiments, analytical work and numerical simulations in which PV homogenisation is observed, suggesting it to be a robust, ubiquitous mechanism. The paper then goes on to report some preliminary results of numerical simulations of decaying quasi-geostrophic turbulence in a channel. As here, they use a CASL model. They find that PV mixing, and therefore the formation of a PV staircase profile where well mixed regions are separated by a sharp PV gradient, to be strongly dependent on initial conditions such as the initial PV gradient and the Rossby radius. In many cases the PV staircasing remains incomplete. However, they present one striking example where the time evolution of the mean y position of the PV contours shows that in some regions the contours collapse together while in other regions they spread apart. Comparing this with the zonally averaged zonal velocity at the final time shows that this inhomogeneous mixing has resulted in strong concentrated eastward jets where the PV contours are close together, separated by broader westward jets where the PV contours are further apart. These jets are a consequence of PV inversion. As shown in Dritschel and McIntyre [2008], inverting a perfect staircase PV profile produces strong, narrow eastwards jets and broader, weaker westward jets. We expect the same to be the case in the shallow water simulations considered here, even though PV inversion depends on the assumption of an underlying balance. We shall see in the following chapter that these flows are indeed close to balanced. Here we apply the techniques used in Dritschel and McIntyre [2008] to our spherical shallow water simulations to investigate the PV homogenisation process in spherical shallow water turbulence.

We begin by calculating $\bar{z} = \overline{\sin\phi}$, i.e. the mean of the sine of the latitudinal position, at each time, of each PV contour that wraps the sphere. Such contours are easily identified as they have a winding number of 2π. In the simpler channel case, the contours that do not wrap the domain represent small scale vortices and

filaments. This is still the case for many of the flows considered here. However, the strong global scale equatorial wave that is present in many of the runs (see section 3.1.2) can cause a precession of the polar vortex that displaces it off the pole resulting in PV contours being classified as wrapping contours at some times but not others. Sometimes the displacement of the polar vortex is so strong that the sharp PV gradient around its edge will not be captured by this analysis.

In figure 3.18, we plot the evolution of $\bar{z}$ for the PV contours that wrap the sphere alongside the initial and final latitudinal profiles of the zonally averaged zonal velocity. We see that there is a significant difference between the cases both in terms of the extent of PV homogenisation and the strength of the jets. Unlike the simulations in Dritschel and McIntyre [2008], we do not start with a zonal velocity profile that is independent of latitude. Our ramped initialisation has resulted in initial states which already contain jets that are then rapidly sharpened as the flow evolves. In general, the results agree with those found in the channel case: regions of bunched up PV contours correspond to eastward jets whereas regions where the PV contours have spread correspond to westward jets. This is most obvious in cases 2 and 4 where the PV is almost entirely homogenised throughout the entire Northern hemisphere. In both these cases just one eastwards jet and one westwards jet are formed. However, these are both cases where the 'polar' vortex is strongly disrupted by the equatorial wave. In case 3 the equatorial wave is weaker and has less of an effect on the polar vortex so in this case we observe the formation of an additional polar jet.

The problem now is how to quantify this over the Fr-Ro parameter space. We do this by employing cluster analysis techniques that are well known in statistical data analysis, particularly in the biological sciences. The idea is to look at the average latitudinal position of the PV contours at time t and deduce whether there are regions where the contours are bunched together, i.e. have formed a cluster. Applying this at each time will show whether the PV contours have become more

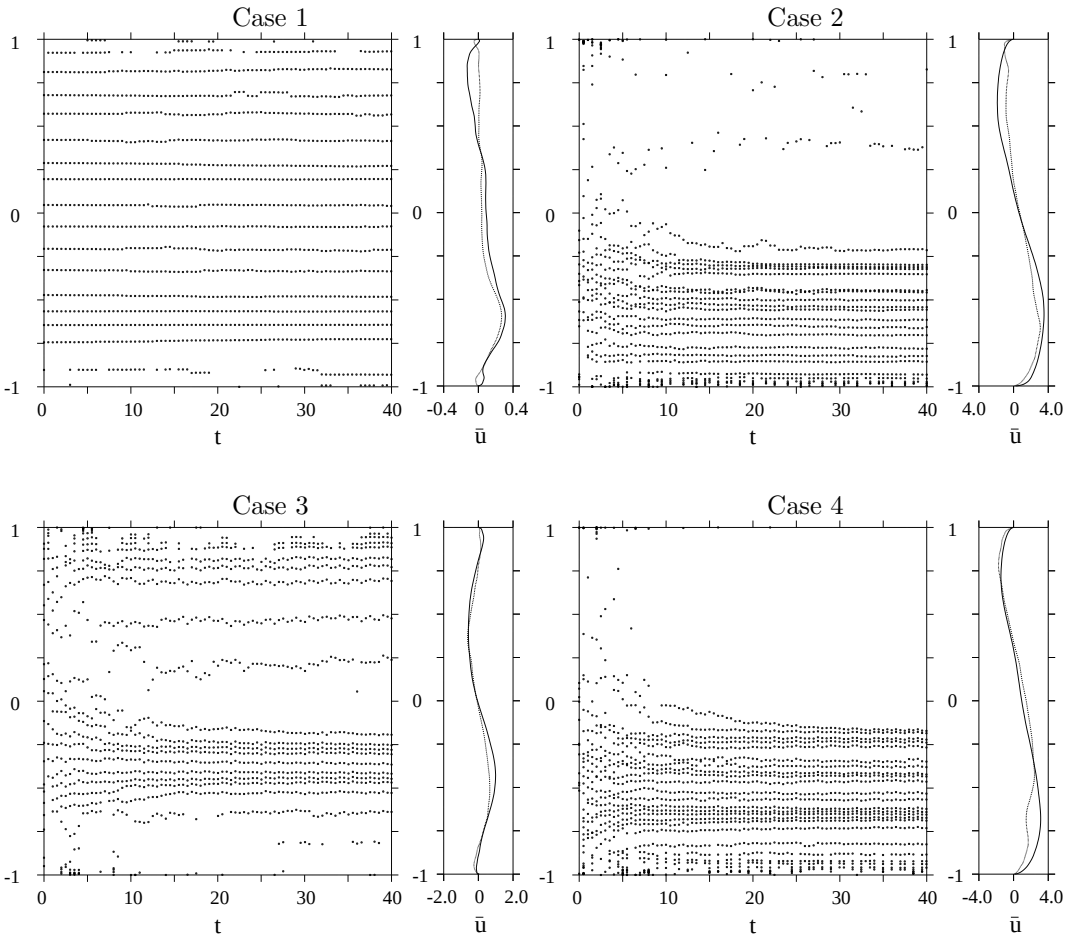


Figure 3.18: For each of the four cases we plot $\bar{z}$, the mean of the sine of the latitudinal position, of each PV contour that wraps the pole, versus time. We also plot the initial (dashed line) and final (solid line) mean zonal velocity profile.

clustered over time. There are many algorithms for partitioning data and evaluating the quality of the resulting clusters. The partitioning step always involves the specification of either the number of clusters or the maximum diameter used to define a cluster. Since, *a priori*, we have no preconceptions about either of these numbers, we must run the algorithm multiple times with different values and find which best represents the data. The arguments concerning speed and effectiveness of clustering algorithms are largely irrelevant here as our data set is small and one dimensional. However, since there may be regions where the PV mixing has been incomplete, we do expect to find some PV contours that do not belong to any cluster. These outliers can strongly affect so-called hierarchical clustering algorithms that build up clusters based on previously established clusters [Jain et al., 1999]. Two popular non-hierarchical algorithms exist: K-means and QT-clust (quality threshold clustering) [Jain et al., 1999, Heyer et al., 1999]. The K-means algorithm requires initial cluster 'seeds' to be specified. Data points are then assigned to the nearest cluster seed. As there is no obvious way to specify these cluster seeds, we prefer the QT-clust algorithm. This algorithm works by calculating all the clusters that satisfy a maximum diameter criterion and then saving the cluster containing the most points. In the event that two or more clusters contain the same number of points, the most coherent cluster (i.e. that with the least variance) is chosen. These points are removed from consideration and the process is repeated until all clusters have been found.

As mentioned above, we have no preconception about the properties of the clusters. Since the QT-clust algorithm requires the maximum diameter, $\mathscr{D}_{max}$, for a cluster to be specified, we have run the algorithm ten times with $\mathscr{D}_{max}$ ranging from π/n_c (where n_c is the number of contours) to π for each time t. The lower limit on $\mathscr{D}_{max}$ would class each PV contour as a cluster if they were evenly spaced; the upper limit takes the other extreme and groups all the contours together as one cluster. Then, again for each t, we compute cluster validity statistics

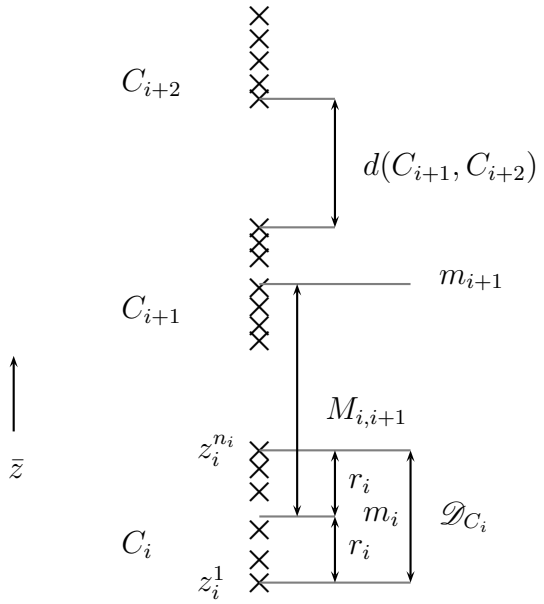


Figure 3.19: Schematic illustrating some of the cluster notation summarised in table 3.3. The crosses show the position of the data at one particular time.

Cluster validity indices	
D	Dunn index
σ_{rms}	Root mean squared standard deviation
Υ	S_Dbw
Cluster notation	
S	total set of points
n	total number of clusters
$1 \leq i \leq n$	index of cluster
C_i	denotes cluster i
$n_i = \|C_i\|$	number of points in C_i
z_i^j	denotes the jth member of cluster C_i
$\bar{z}_i = \frac{1}{n_i}\Sigma_{j=1}^{n_i} z_i^j$	mean value of cluster points in C_i
$\bar{z} = \frac{1}{\|S\|}\Sigma_{k=1}^{\|S\|} z_k$	mean value of entire dataset S, where $\|S\|$ is the number of data points in S and $z_K \in S$.
$m_i = \frac{1}{2}(z_i^{n_i} + z_i^1)$	midpoint of cluster C_i
$M_{i,i+1} = \frac{1}{2}(m_{i+1} + m_i)$	midpoint between midpoints of clusters C_i and C_{i+1}
Distance notation	
$d(p_1, p_2) = \|p_1 - p_2\|$	distance between points p_1 and p_2
$\mathscr{D}_{C_i}$	diameter of cluster C_i
$r_i = \frac{1}{2}\mathscr{D}_{C_i}$	radius of cluster C_i
$\mathscr{D}_{max}$	maximum cluster diameter (required as input for the QT-clust algorithm)
Statistical notation	
$\sigma(S)$	variance of dataset
$\sigma(C_i)$	variance of cluster
ς	intra-cluster scattering
$\varrho(p)$	density of data points about point p, i.e. the number of data points within one standard deviation of point p

Table 3.3: A summary of cluster notation.

to determine which $\mathscr{D}_{max}$ produces the most coherent, well separated clusters. Many cluster validity indices have been defined [see Jain et al., 1999, Halkidi et al., 2001, Kovàcs et al., 2005, for a comprehensive review]. Before applying this method to all our simulations, we have looked in detail at the results given by four different validity statistics for our four cases. The validity statistics are defined below for 1 dimensional data, partitioned into n clusters with C_i denoting cluster i and $n_i = |C_i|$, i.e. the number of points in C_i. The clusters are ordered from South to North, as are the points within each cluster (see figure 3.19 for a schematic and table 3.3 for a summary of the notation).

- The Dunn index

$$D = \min_{i=1,n} \left(\min_{j=i+1,n} \left(\frac{d(C_i, C_j)}{\max_{k=1,n}(\mathscr{D}_{C_k})} \right) \right), \tag{3.5}$$

where

$$d(C_i, C_j) = \min_{z \in C_i, z' \in C_j} |z - z'|, \tag{3.6}$$

is the minimum distance between cluster C_i and cluster C_j and

$$\begin{aligned} \mathscr{D}_{C_i} &= \max_{z,z' \in C_i} (d(z, z')) && (3.7) \\ &= d(z_i^{n_i} - z_i^1) && (3.8) \\ &= |z_i^{n_i} - z_i^1| && (3.9) \end{aligned}$$

is the diameter of cluster C_i. Note that $\max_{k=1,n}(\mathscr{D}_{C_k}) \neq \mathscr{D}_{max}$ since the first is the maximum diameter of the clusters produced whereas the second is the maximum permitted cluster diameter that is specified before the clustering algorithm is run.

For well separated clusters, the distance between clusters, $d(C_i, C_j)$, will be large; for coherent clusters the diameter of individual clusters, $\mathscr{D}_{C_i}$, will be small. Therefore we expect a large value of D to imply that we have found a good partition of the data.

- Root mean squared standard deviation

$$\sigma_{rms} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{n_i}(z_i^j - \bar{z}_i)^2}{\sum_{i=1}^{n}(n_i - 1)}}, \tag{3.10}$$

where

$$\bar{z}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} z_i^j$$

is the mean $z = \sin\phi$ position of all points in cluster C_i. This uses the variance of the clusters as a measure of their coherence. Coherent clusters will have a small variance so we expect a small value of σ_{rms} to indicate that the clusters are good. Note that this measure does not tell us anything about how well separated the clusters are.

- S_Dbw

This is the most sophisticated index included here as, in addition to measuring cluster coherence and separation, it takes into account the density of the clusters. It is defined in Halkidi and Vazirgiannis [2001] (using different notation) as

$$\Upsilon = \varsigma + \bar{\varrho}, \tag{3.11}$$

where ς represents the intra-cluster scattering (compactness of the clusters) and $\bar{\varrho}$ represents the inter-cluster density. The intra-cluster scattering is based on the variance of each cluster, compared to the variance of the original dataset. The variance within each cluster is defined as:

$$\sigma(C_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} (z_i^j - m_i)^2, \tag{3.12}$$

where m_i is the midpoint of cluster C_i. The variance of the dataset S is:

$$\sigma(S) = \frac{1}{|S|} \sum_{k=1}^{|S|} (z_k - \bar{z})^2, \tag{3.13}$$

where $|S|$ is the number of elements in S and $\bar{z}$ is the mean value of S. Now we can define the intra-cluster scattering measure:

$$\varsigma = \frac{1}{n} \sum_{i=1}^{n} \frac{\sigma(C_i)}{\sigma(S)}. \tag{3.14}$$

Coherent clusters will have a small variance compared with the variance of the data so a small value for ς implies that the clusters are well defined. In order to check whether they are well separated we compute the inter-cluster density:

$$\bar{\varrho} = \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \frac{\varrho(M_{i,i+1})}{\max(\varrho(m_i), \varrho(m_{i+1}))}, \tag{3.15}$$

where $M_{i,i+1}$ is the midpoint between the midpoints of clusters i and $i+1$ and the $\varrho(p)$ function counts the number of elements within an average standard deviation of the point p. The average standard deviation is simply $1/n\sqrt{\sum_{i=1}^{n} \sigma(m_i)}$. For a dense cluster C_i, $\varrho(m_i)$ will be large and for well separated clusters i and $i+1$, $\varrho(M_{i,i+1})$ will be small. So, combining ς and $\bar{\varrho}$, we expect a low value of Υ to imply that we have found a good partition of the data.

To summarise the method, for each time t:

1. Calculate average latitudinal position of each PV contour that wraps the pole.

2. Run the QTclust algorithm:

   (a) Specify a maximum diameter $= \dfrac{2\pi}{n_c}$.

   (b) Calculate all possible clusters with the maximum diameter.

   (c) Save the cluster containing the most points and remove these points from consideration.

   (d) Repeat steps 2b-2c on the remaining points until there are no more possible clusters. Note that a single point does not count as a cluster.

3. Repeat the QTclust algorithm, increasing the maximum diameter.

4. Calculate the validity indices outlined above for each cluster partition produced.

5. For each validity index, choose the set of clusters it deems best.

Figures 3.20-3.21 show the best clusters, according to the three validity indices, for each of our four cases. For each time, each cluster is plotted with a different symbol. Unfortunately, since there is no way to link the clusters from one time to the next, the symbols do not bear any relation to each other at different times. Reassuringly, all three validity indices pick out almost identical clusters over the last half of the simulations, i.e. once the flows have adjusted. Case three is the exception. Although the Dunn index and S_Dbw validity indices agree, the r.m.s. standard deviation picks out far more structure, giving 4-5 clusters compared to the 2 clusters picked out by the other methods. For ease of comparison, we plot the clusters selected by each validity index at $t = 22$ next

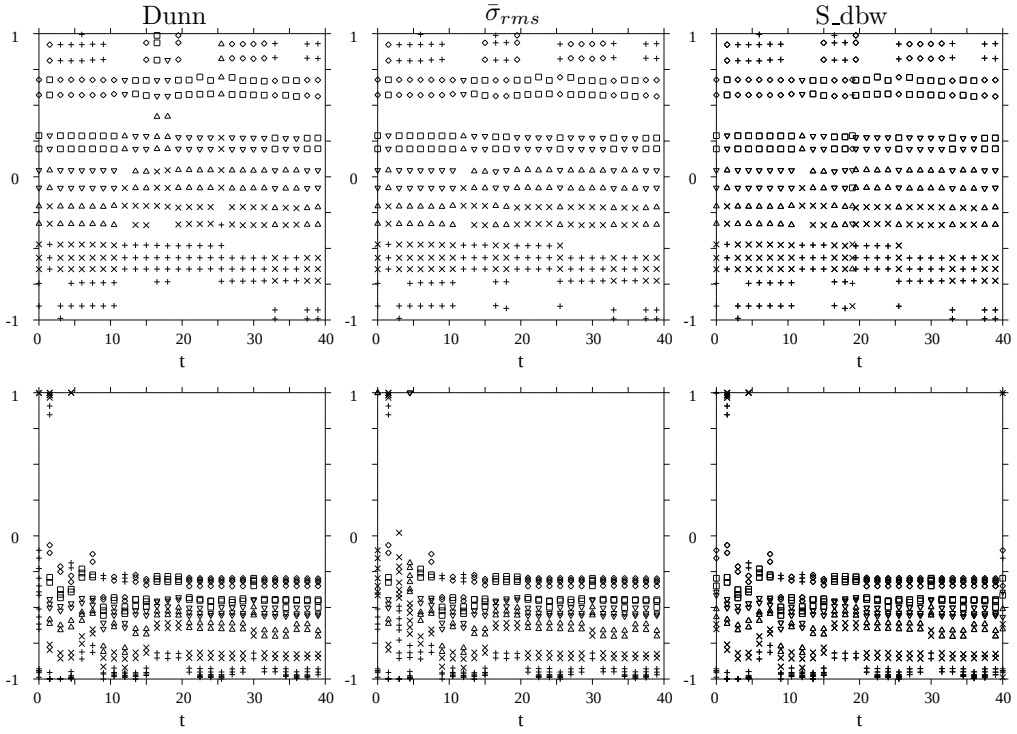


Figure 3.20: The 'best' partitions of the data into clusters according to the Dunn index (left column), the r.m.s. standard deviation (middle column) and S_Dbw (right column) for cases 1 (top row) and 2 (bottom row). The y-axis is $\bar{z}$ from -1 (south pole) to +1 (north pole) and time runs from 0-40 days along the x-axis. Within each time level, each cluster is plotted using a different symbol. Note that, since there is no way to link the clusters from one time to the next, these symbols do not bear any relation to each other at different times. These plots contain fewer data points than the corresponding plots in figure 3.18 as 'clusters' containing just one data point have been removed.

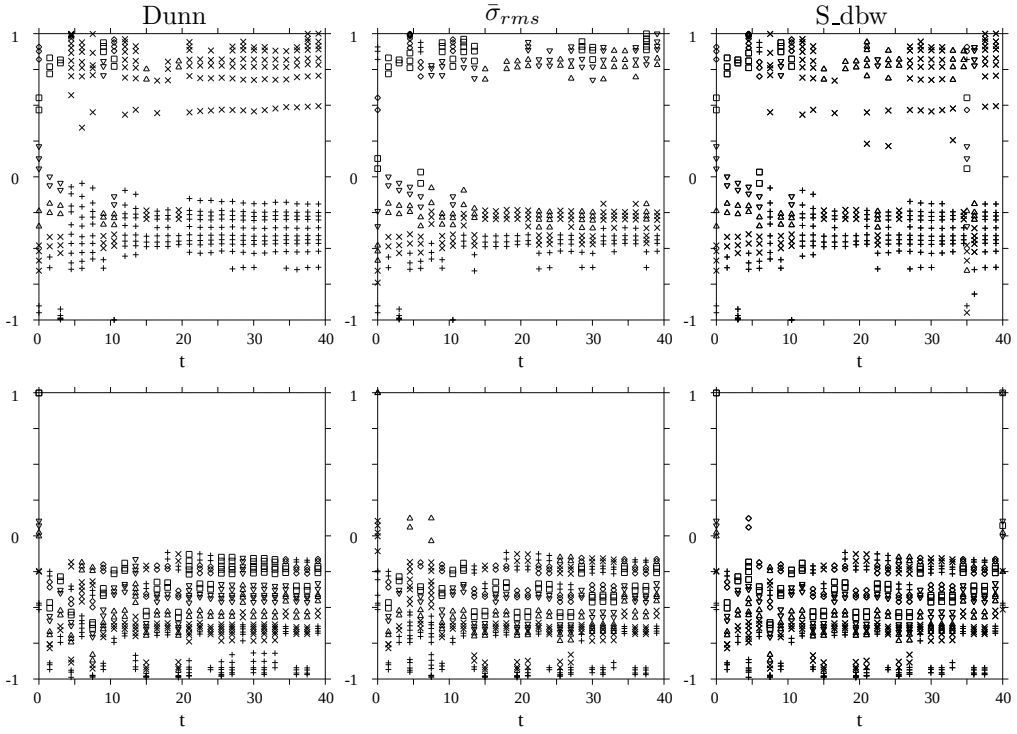


Figure 3.21: The 'best' partitions of the data into clusters according to the Dunn index (left column), the r.m.s. standard deviation (middle column) and S_Dbw (right column) for cases 3 (top row) and 4 (bottom row). The y-axis is $\bar{z}$ from -1 (south pole) to +1 (north pole) and time runs from 0-40 days along the x-axis. Within each time level, each cluster is plotted using a different symbol. Note that, since there is no way to link the clusters from one time to the next, these symbols do not bear any relation to each other at different times. These plots contain fewer data points than the corresponding plots in figure 3.18 as 'clusters' containing just one data point have been removed.

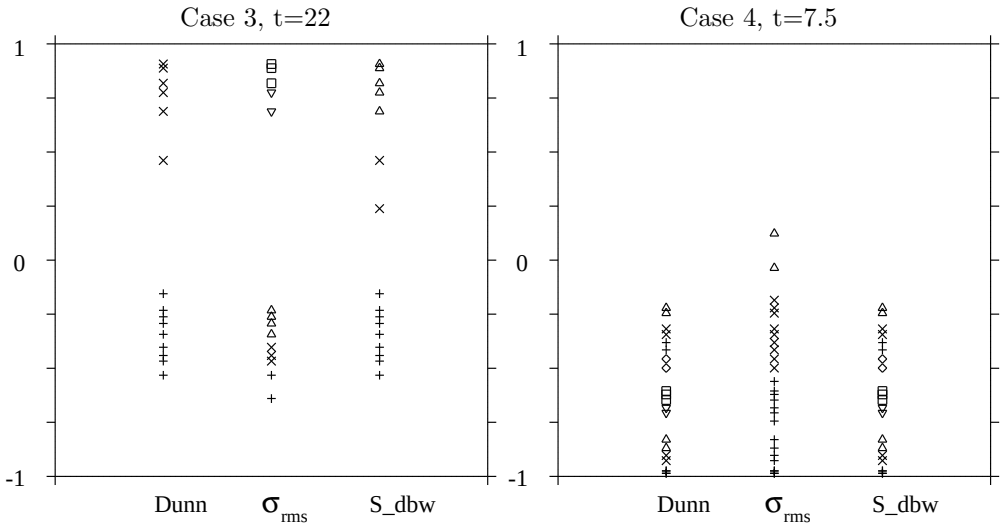


Figure 3.22: Two examples contrasting the types of clusters preferred by each of the three validity indices. The example on the left is from case 3 at t=22. The example on the right is from case 4 at t=7.5.

to each other in the left hand plot of figure 3.22. At this time, S_Dbw picks out 3 clusters whereas the r.m.s. standard deviation picks out 5, despite rejecting one of the clusters deemed satisfactory by S_Dbw. The Dunn index only sees two large, sprawling clusters. It would seem that the r.m.s. standard deviation prefers clusters that are compact, but not necessarily well populated, whereas the Dunn index and S_Dbw prefer the clusters to be denser but more separated. This is consistent with the way the validity indices are defined. Unfortunately, it is possible to find occasions where the opposite occurs. The right hand plot of figure 3.22 shows such an example, occurring in case 4 at $t = 7.5$. This time it is the Dunn and S_Dbw indices that pick out the most structure finding 9 clusters compared to the 3 selected by r.m.s. standard deviation. However, this example occurs during the the adjustment stage and for the second half of the simulation, r.m.s. standard deviation again becomes the measure that sees the most structure.

Case	Dunn index	σ_{rms}	S_Dbw
(1)	0.847	0.076	0.010
(2)	0.837	0.023	0.025
(3)	1.463	0.047	0.053
(4)	0.581	0.022	0.009

Table 3.4: Average values of the three validity indexes over the last 10 days of the four simulations. Remember that a *large* value of the Dunn index indicates that the data is well clustered whereas for σ_{rms} and S_Dbw this is indicated by *small* values.

Figures 3.20 and 3.22 have demonstrated that the properties of the clusters deemed best by each validity index vary. A subtlety to bear in mind is that we have selected the *best possible* cluster partition; it could be that the contours are not clustering at all, or that they are clustering more in one case than in another. To see this, we have to look in more detail at the values of the validity indices. These are presented in table 3.4, averaged over the last 10 days of the simulation. These values indicate a surprising result for two of the three validity indices: the Dunn index and S_Dbw consider the PV contours in cases 3 and 1 respectively to be the most clustered whereas, by eye, we can see that the PV contours in case 1 and, to a lesser extent, case 3 are hardly clustered at all! The r.m.s. standard deviation correlates much better with our visual observations.

Returning to our objective of quantifying the jets over the Fr-Ro parameter space, we examine the clustering results for all of the simulations. As we have seen above, the number of jets predicted by each validity index is erratic due both to the precession of the 'polar' vortices and to the nature of the clustering algorithm which is, despite all our efforts, very sensitive. A more robust measure appears to be the average value of the validity index over the last 10 days of each simulation. This number tells us if the validity index perceives the PV contours to be well

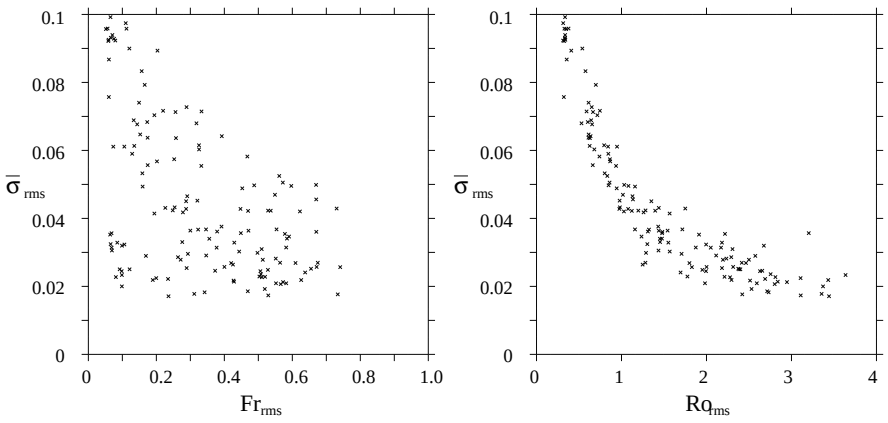


Figure 3.23: $\bar{\sigma}_{rms}$ versus Fr (left hand figure) and versus Ro (right hand figure) Remember that *small* values of $\bar{\sigma}_{rms}$ indicate well clustered contours.

clustered or not. Unsurprisingly, given the results in table 3.4, the r.m.s. standard deviation gives the most informative results. Figure 3.23 shows $\bar{\sigma}_{rms}$ versus Fr_{rms} (on the left) and versus Ro_{rms} (on the right). We see that there is a slight correlation between $\bar{\sigma}_{rms}$ and Fr, with the PV contours becoming more clustered with increased Froude number. However, the correlation between $\bar{\sigma}_{rms}$ and Ro is striking. $\bar{\sigma}_{rms}$ decreases with increasing Ro, levelling out at around Ro = 2. This suggests that increasing the importance of rotation (i.e. smaller Rossby number) inhibits the clustering of PV contours. This seems counterintuitive and is contrary to predictions based on the Rhines scale L_{Rh} (defined in equation 3.18 below), but Cho and Polvani [1996] find that a similar result emerges from their spherical simulations. They suggest that the variation of β with latitude coupled with the lack of forcing, causes the jets to be ill defined. They note that the only obvious strong jets are those surrounding the polar vortices and, as mentioned above, the polar vortices pose a problem for our method since they rarely sit neatly over the pole.

Two factors seriously inhibit the performance of the clustering approach: the constraint the PV contours must wrap the sphere and the longitudinal averaging.

The first causes the strong jets around the 'polar' vortices to be neglected, the second smears out the complex longitudinal structure of the jets. We have found that a local analysis of the jet characteristics is much more revealing. This can be achieved by examining the palinstrophy field

$$P = \frac{1}{2}|\nabla \Pi|^2, \tag{3.16}$$

which measures the strength of the gradient of the PV. A large palinstrophy value will indicate local tightening of PV contours which is exactly what we expect to find in westerly jets. Figures 3.24-3.27 show the time evolution of P for the four cases. We plot the scaled palinstrophy ($\log_{10}(P/(3 \times P_{rms})$) in order to highlight the areas with the steepest gradients. Also plotted is the spatial distribution of kinetic energy. Note that the times shown are $t = 0, 5, 10, 20$ and 40 so that the first 4 figures show the early evolution during which the adjustment is occurring and the final figure shows the long term behaviour.

Structures in the plots of kinetic energy and the palinstrophy field are closely correlated although, since the kinetic energy depends on the global distribution of PV, it is smoother and does not reveal the smaller scale features evident in the palinstrophy plots. Since the jets meander, it may be that a significant component of the local flow is in the cross-jet direction and this may mask the presence of strong flow along the jet axis in the kinetic energy plots. The palinstrophy plots suffer from no such difficulties. The detailed structure evident in these plots enables us to precisely locate the jet axes as the regions where the PV gradient is strongest (see the dark regions in the palinstrophy plots, figures 3.24-3.27). We can see immediately from figure 3.24 that the constraint imposed for the cluster analysis, that all jets must wrap the sphere, is too harsh. Particularly in the earlier plots, we see jets that persist for only one quarter of the latitude band. The last frame of figure 3.27 demonstrates the complex, meandering paths that the jets take around the sphere. Even though, by this stage, the jets persist over

greater zonal distances, the latitudinal location of the jet axis varies so much with longitude that the averaging process necessary for performing the cluster analysis will obscure the presence of the jet. Note that the jet visible in the northern latitudes of the last frame of figure 3.24 is aligned almost north-south at zero longitude.

Cases 2-4 (figures 3.25-3.27) illustrate another feature of the jets that again proves a problem for the cluster analysis. In these three cases the final frames indicate that the flow evolves into a final state that consists of strong jets near the equator and strong 'polar' vortices. These vortices are not strictly *polar* vortices since they precess (to differing degrees in each case) off the pole. In all three cases the precession is so strong that the vortex is sufficiently displaced that the clustering algorithm is unable to see the jets surrounding it as they no longer wrap the pole. The structure of these vortex-surrounding jets can be highly complex, as illustrated by frames 2-5 of figure 3.27.

In case 2 we note that the jets around the equator are more zonal and meander less than those in cases 3 and 4, disturbed only when the large northern hemisphere vortex passes. In case 3 the strong vortex remaining in the northern hemisphere stays much closer to the pole. Case 4 demonstrates the opposite effect to that noted by Youssef and Marcus [2003] who commented that the strong zonal jets on Jupiter enabled the merger of vortices that would never have remained close for long enough to merge without the constraint placed on their motion by the jet. In this case we see that the interaction between the jets and the vortices produces strongly meandering jets that then inhibit the merger of the two remaining northern hemisphere vortices which are trapped in the meanders of the jet.

Noticing the strong correlation between features in the palinstrophy and kinetic energy plots, we use this to diagnose a jet velocity

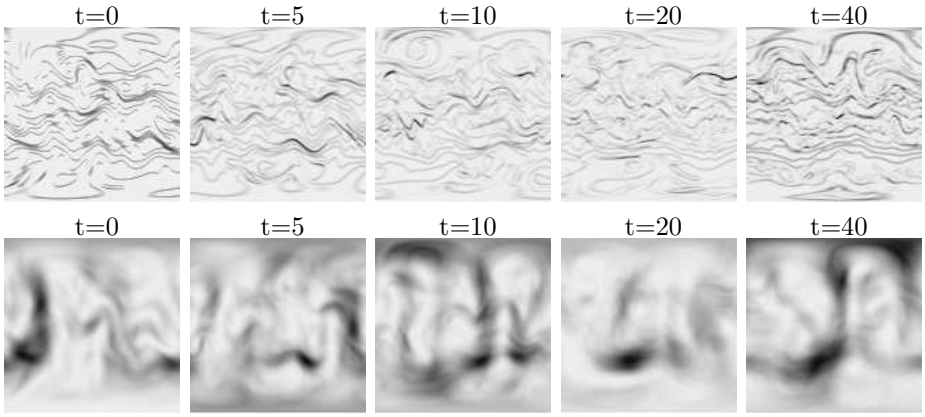


Figure 3.24: Evolution of palinstrophy (top) and kinetic energy (bottom) for case 1. For each figure the range has been determined to optimise the detail shown in the figure so these figures are for qualitative rather than quantitative comparison. Note also that the figures are not evenly spaced in time.

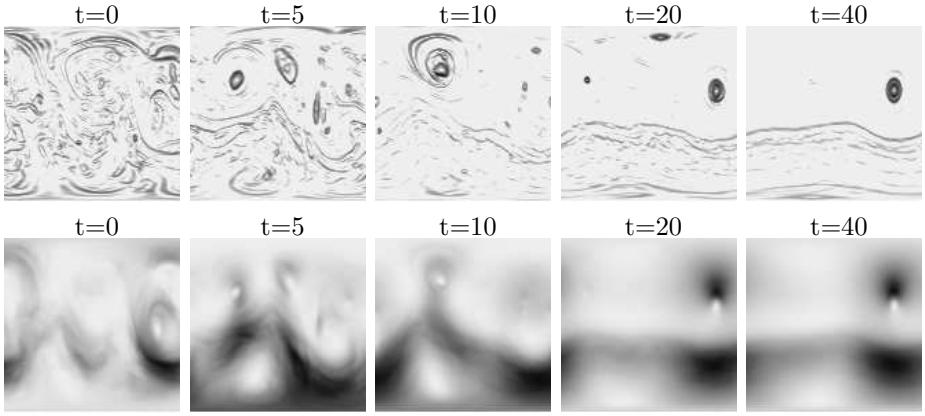


Figure 3.25: Evolution of palinstrophy (top) and kinetic energy (bottom) for case 2. For each figure the range has been determined to optimise the detail shown in the figure so these figures are for qualitative rather than quantitative comparison. Note also that the figures are not evenly spaced in time.

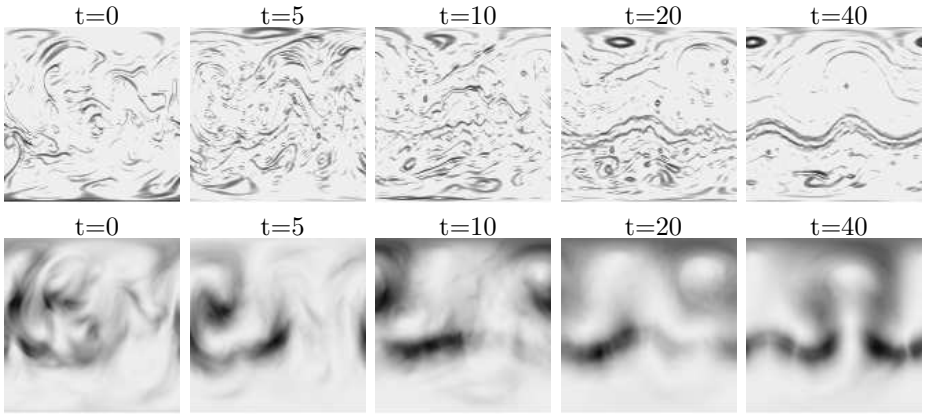


Figure 3.26: Evolution of palinstrophy (top) and kinetic energy (bottom) for case 3. For each figure the range has been determined to optimise the detail shown in the figure so these figures are for qualitative rather than quantitative comparison. Note also that the figures are not evenly spaced in time.

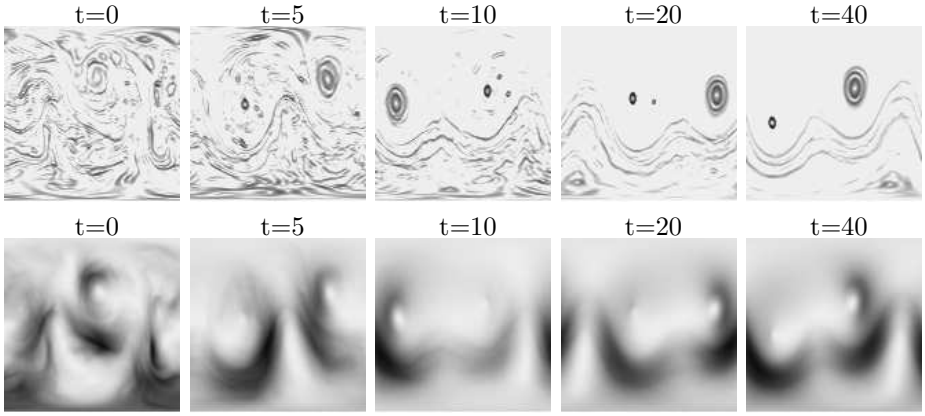


Figure 3.27: Evolution of palinstrophy (top) and kinetic energy (bottom) for case 4. For each figure the range has been determined to optimise the detail shown in the figure so these figures are for qualitative rather than quantitative comparison. Note also that the figures are not evenly spaced in time.

$$U_{jet}(\phi) = \sqrt{\left(\frac{< P(u^2 + v^2) >}{< P >} \right)}, \tag{3.17}$$

where $< . >$ now denotes the average over longitude. We now calculate two measures of the local jet spacing: the Rhines scale

$$L_{Rh}(\phi) = \sqrt{\left(\frac{U_{jet}(\phi)}{\beta(\phi)} \right)} \tag{3.18}$$

and the jet spacing b defined in Dritschel and McIntyre [2008] and obtained by inverting, at each latitude, their equation (6.3)

$$\left(\frac{L_{Rh}}{L_D} \right)^2 = \frac{b}{L_D} \tanh \left(\frac{b}{2L_D} \right), \tag{3.19}$$

where L_{Rh} and L_D are now functions of latitude. These two measures are plotted against latitude in figure 3.28 for each of the four cases. We see that b always gives a larger jet spacing than L_{Rh} although both are of the same order of magnitude. Both measures blow up towards the poles in line with the theory that there is a critical latitude above which jets will not form [Theiss, 2004, Showman, 2007]. Cases 1 and 3 have smaller jet spacings, indicating a greater number of jets, than cases 2 and 4. We see from figure 3.24 that case 1 does indeed exhibit many jet 'streaks', i.e. jets that do not wrap the sphere. These features were not picked up by the clustering algorithms. The final state of case 3 (figure 3.26) also contains some jet streaks although most of the strong jet activity is concentrated in the northern hemisphere vortex and two jets just south of the equator. These two jets cause a large bump in the jet spacing. The jet spacing plots for cases 2 and 4 show a marked difference between the northern and southern hemispheres. In the northern hemisphere the jet spacing is noisy due to the strong, coherent vortices present (see figures 3.25 and 3.27) whereas for most of the southern hemisphere

the jet spacing increases linearly away from the equator. Unfortunately there are only two jets present in each of these simulations so it is not possible to see this widening of jet spacing in the palinstrophy plots.

This new measure of jet spacing has yielded some intriguing insights that could be followed up by examining flows that could be expected to generate more jets. Another possibility is to return to the PV contours and calculate an along-jet velocity in regions where the contours are tightly spaced. However, recent (unpublished) work by Henderson and Dritschel suggests that this approach does not yield significantly different results from the simpler approach taken here.

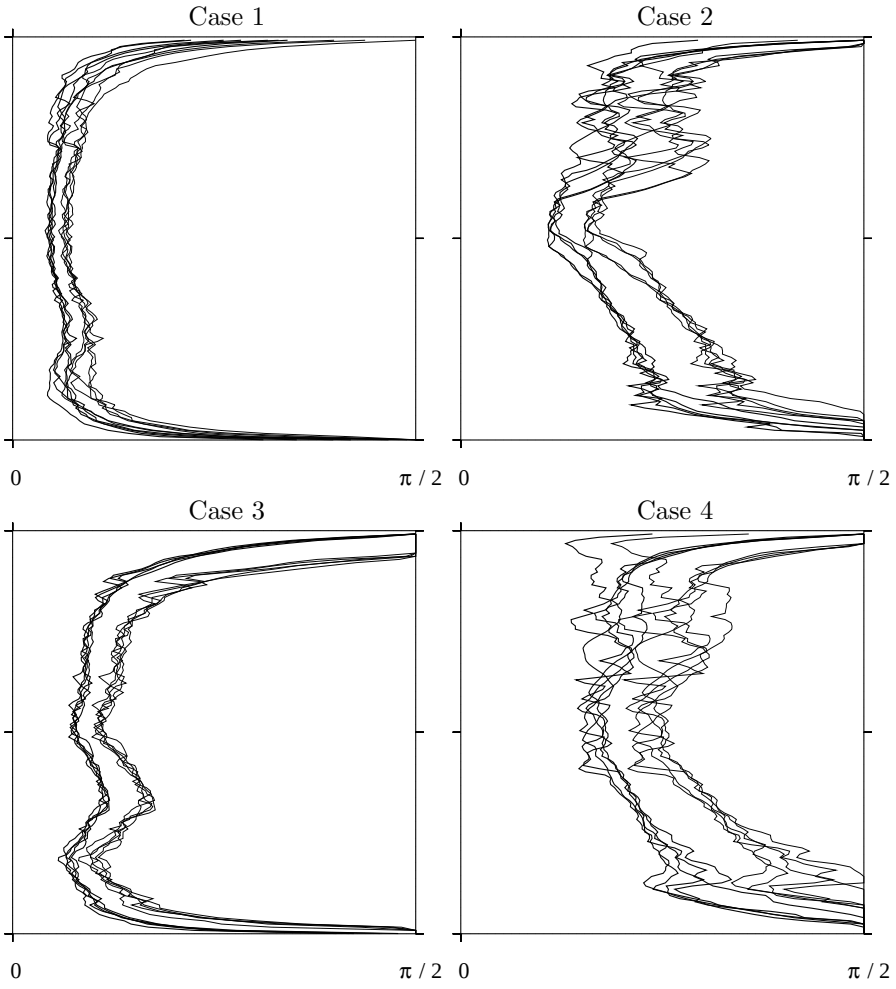


Figure 3.28: Two measures of jet spacing: L_{Rh} (bold line) and b (thin line), plotted against latitude (y axis) for each of the four cases at $t = 20, 25, 30, 35$ and 40.

Chapter 4

Wave–vortex decomposition

In this chapter we examine the extent to which the turbulent flows of chapter 3 can be considered balanced. We repeat the 150 simulations using two balanced models and compare the results with the balanced components of the original flows, computed using a novel wave-vortex decomposition algorithm.

4.1 Definitions of balance

We saw in chapter 2 that there is a distinct separation between vortical (balanced) and wave (unbalanced) motion in the shallow water equations if they are linearised about a state of rest. Defining a balanced flow is not so straightforward in the nonlinear case and there are many different approaches. One approach is to create a balanced model by restricting the model equations to explicitly exclude inertia-gravity waves. The quasi-geostrophic model presented in section 2.3 is an example of this type of model but more accurate models can be derived by retaining higher order terms. While useful for exploring the qualitative properties of geophysical flows, these models constrain the flow to remain in the region of parameter space for which they are valid, i.e. the equations impose a balance

that might not be there if the flow were permitted to evolve according to the full equations. *A balanced flow produced in this way will not necessarily be a solution of the full fluid equations.* Indeed, vortical flows are known to spontaneously emit inertia-gravity waves [Williams et al., 2003, Viúdez and Dritschel, 2006, and references therein] so a balanced flow which is *also* a solution of the full fluid equations must nevertheless contain a small *minimal* amount of wave motion.

A second approach is to filter the output of an inertia-gravity wave permitting model to remove the waves but retain the 'balanced' component. This approach makes use of the assumption that the frequency separation explicit in the linearised equations carries over, at least to some extent, to the nonlinear equations. We shall see in section 4.3 that this is not necessarily the case. Indeed, it is worth noting that the assumption of a clear frequency separation is not generally true even for the linearised equations. An interesting example is that of a mean flow containing a PV jump [Dritschel and Vanneste, 2006].

Recent research [Viúdez and Dritschel, 2004b] has exploited a third approach: obtaining the balanced flow directly from the inertia-gravity wave permitting model. This technique exploits the properties of PV to produce an 'optimally' balanced flow in the sense that minimal wave motion, i.e. that due to spontaneous generation, is present. This novel wave–vortex decomposition algorithm will be described in more detail in section 4.1.2.

4.1.1 The $\delta - \gamma$ hierarchy

Balanced models consist of one prognostic equation and a set of diagnostic balance relations that determine the remaining variables. The most common balanced model used for geophysical flows is the quasi-geostrophic model outlined in section 2.3. In section 2.3 this model was derived by performing a scale analysis on the non-dimensional shallow water equations and retaining just the $O(\mathrm{Ro})$

terms (with Fr $\sim O(\mathrm{Ro})$). This yields a prognostic equation for the depth field which, coupled with the geostrophic relations 2.38 and 2.39, defines a balanced model. However, we have stressed the central role of the PV field as the field which determines the balanced motion. It is possible to define a 'QG PV' (see 2.3) but this involves approximation of the very field that we wish to accurately represent in order to correctly model the balanced motion. An alternative approach taken by McIntyre and Norton [1999], McKiver and Dritschel [2008] and Mohebalhojeh and Dritschel [2000a] is to keep the form of, and the prognostic equation for, the unapproximated PV and explore the possible ways to diagnose the remaining variables. Mohebalhojeh and Dritschel [2000b] provide a thorough exploration of the possible hierarchies based on the δ and γ variables introduced in section 2.2. The δ and γ variables are, not coincidentally, the same variables that are used to represent the leading order unbalanced motion in the CASL model described in sections 2.5-2.6.

The setup of the CASL model is such that it is relatively straightforward to run simulations using the first two members of the δ–γ hierarchy. Although originally developed as a simplification of a formal multiple-timescale Rossby number expansion (the Baer-Tribbia nonlinear normal-mode initialisation method Baer and Tribbia [1997]), the δ–γ hierarchy is not based on a formal asymptotic scaling. Instead it can be justified via an expansion of the nonlinear terms in the normal mode representation of the shallow water equations [Mohebalhojeh and Dritschel, 2000b]. Some members of this hierarchy have been used previously (see refs in Mohebalhojeh and Dritschel [2000a]) but this is the first time that they have been applied on a full sphere. Mohebalhojeh and Dritschel [2000a] report that the hierarchies of balance conditions are asymptotic: as the order increases the imbalance decreases but then increases again. For the f-plane shallow water equations they find the fourth member of the hierarchy (in which the third derivatives of δ and γ are set to zero) to be the most accurate. However, for

the spherical case we find (see section 4.2) that even the second member of the hierarchy (in which δ_t and γ_t are set to zero) frequently fails to converge for the higher Froude number cases. Hence we do not proceed with any of the higher members. The results of Smith and Dritschel [2006] suggest that even zeroth order δ–γ balance (in which δ and γ are set to zero) offers significant improvements over QG balance, especially at large scales and near the equator. Note, this balance does *not* approximate the form of PV - the exact definition is used.

4.1.2 Optimal potential vorticity balance

The optimal potential vorticity (OPV) balance method of Viúdez and Dritschel [2004b] exploits the property that PV cannot be transferred into inertia-gravity waves. The procedure is closely related to the initialisation procedure, first outlined in Viúdez and Dritschel [2004a] and described in section 3.1.1 in which the PV anomaly is set to zero and smoothly ramped up to the required value while the full equations are integrated over a fictitious time period. It is possible to specify that the PV contours will remain fixed, as in the dynamical PV initialisation procedure of Viúdez and Dritschel [2004a], the idea being that the other fields will slowly adjust to this required PV distribution. However, *this restriction on the PV evolution may result in a flow that is no longer a solution of the full equations.* Instead it is preferable to allow the PV contours to evolve. Starting from the flow configuration at time t, the aim is to iterate over a fictitious time τ, removing inertia-gravity waves with each iteration. In the spherical shallow water case this is accomplished by setting δ and γ to zero at the end of each backward iteration. At the end of each forward iteration a small amount of imbalance will have been generated, so δ and γ will no longer be exactly zero, but the flow will be closer to a balanced state. The PV contours will also be slightly different from the true PV contours of the flow. However, this difference is small and the PV contours can be reset to their original values. The process is repeated until the

difference between successive flow states is sufficiently small. The procedure is summarised below:

1. Save all fields at time t.

2. Fix the PV contours and set the PV anomaly, $\varpi = \Pi - f$, and all other fields to zero.

3. Integrate the model from $\tau = 0$ to $\tau = \Delta\tau$, multiplying ϖ by the ramp function $T(\tau)$ to obtain an initial balanced guess for the other fields.

4. Integrate the model backwards from $\tau = \Delta\tau$ to $\tau = 0$, again except for ϖ which is multiplied by the ramp function $T(\tau)$ but this time (and henceforth) allowing the PV contours to move.

5. Set $\delta = \gamma = 0$.

6. Integrate the full shallow water model forwards from $\tau = 0$ to $\tau = \Delta\tau$, multiplying ϖ by the ramp function $T(\tau)$.

7. Restore the PV contours to their original positions at time t, leaving all other fields unchanged. Save all fields to check convergence at the end of the next forward iteration.

8. Repeat steps 4 to 7, checking for convergence at the end of each forward iteration by comparing the $\tilde{h}$ field to that produced by the previous forward iteration.

We first assess the dependence of the OPV balance output on the diagnostic period $\Delta\tau_D$, the single parameter which needs to be specified in the OPV balance procedure. The diagnosis of balance was carried out using $\Delta\tau_D = 0.5$, 1, 2, 3 and 4 days. Typically, 6-10 forward and backward integration cycles were required for convergence of the procedure (here when the maximum pointwise difference

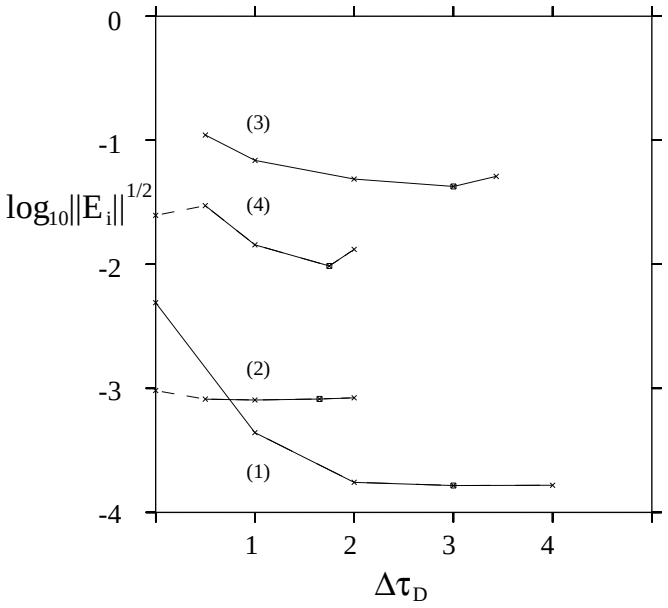


Figure 4.1: Unbalanced energy norm $\left(\langle u_{\mathsf{i}}^2 + v_{\mathsf{i}}^2 + c^2\tilde{h}_{\mathsf{i}}\rangle/2E\right)^{1/2}$ as a function of the ramp period $\Delta\tau_D$ for the four cases described in table 3.1.

in $\tilde{h}$ at the end of successive cycles is less than 10^{-7}). Each case was analysed at 0.5 day intervals from $t = 0$ to 5 (11 samples). This produces a balanced flow, indicated by a subscript b, for each $\Delta\tau_D$. The imbalance, indicated by a subscript i, is just the difference between the actual and the balanced flows, e.g. $\delta_{\mathsf{i}} = \delta - \delta_{\mathsf{b}}$. This also depends on $\Delta\tau_D$. The effectiveness of the wave–vortex decomposition procedure may be judged by the degree to which the flow can be attributed to the PV alone. This is equivalent to minimising the imbalance.

The magnitude of the imbalance is measured by the energy norm

$$||E_{\mathsf{i}}|| = \frac{\langle u_{\mathsf{i}}^2 + v_{\mathsf{i}}^2 + c^2\tilde{h}_{\mathsf{i}}\rangle}{2E}. \tag{4.1}$$

The log (base 10) of the square root of this measure is plotted in figure 4.1 as a function of the ramp period $\Delta\tau_D$ for the four simulations taken from the corners

Case	OPV	OPV (fixed contours)	$\delta_t = \gamma_t = 0$	$\delta = \gamma = 0$
(1)	0.0002	0.0049	0.0049	0.0265
(2)	0.0007	0.0010	0.0010	0.0273
(3)	0.0512	0.0931		0.1890
(4)	0.0142	0.0248	0.0248	0.1450

Table 4.1: The unbalanced energy norm calculated over the first 10 days of the simulation using the four different balancing routines.

of our parameter space (see figure 3.3). Included also are the results for the first-order δ–γ balance ($\delta_t = \gamma_t = 0$) plotted along $\Delta\tau_D = 0$ (when available). Not plotted are the results for cases 2, 3 and 4 using $\Delta\tau_D = 3$ and 4 respectively because these longer ramp times cause the OPV scheme to diverge. For cases 2, 3 and 4 there is an optimal ramp period below which the imbalance has not been minimised and above which the balancing routine will not converge. This is due to the complexity of the PV contours which become tightly wrapped up in these cases. The choice of $\Delta\tau_D$ is therefore dictated by the complexity of the flow. The more complex the PV contours, the shorter the possible ramp period or else the inherent complexity of the flow will cause the iterative procedure to diverge. As a general rule, flows with higher Froude or, to a lesser extent, Rossby numbers will be more complex. Combining this theory with results from initial experiments, leads to the following choice of $\Delta\tau_D$

$$\Delta\tau_D = \min\left(\frac{6}{\mathrm{Ro}}, \frac{2}{\mathrm{Fr}}, 3\right). \tag{4.2}$$

The value of $\Delta\tau_D$ given by equation 4.2 is plotted as squares in figure 4.1 for each of the four cases. This value has been used to produce the results in this section. Many of the results in the following section, however, were computed

using a slightly different version of 4.2 that limited $\Delta\tau_D$ to 5 days rather than 3. Fortunately this does not give significantly different results since the procedure either fails to converge, in which case we simply rerun with a reduced $\Delta\tau_D$, or it converges but takes longer. There is some ambiguity in equation 4.2 concerning the definition of Fr and Ro. In our calculations we have used the initial Froude number and the time mean (over $t = 0 - 10$) rms average Rossby number over the domain. However, as the complexity of these turbulent flows can change dramatically over time, typically peaking in the first few days, it is possible to use a different value of $\Delta\tau_D$ based on the maximum Froude and Rossby numbers reached at time t_{bal}. This could easily be incorporated into the OPV routine.

To avoid this problem entirely it is possible to perform the same iterative procedure but with the PV contours held fixed while the magnitude of the PV jump across each contour is ramped. However, the results are significantly different, with less of the flow attributed to balance. Table 4.1 gives the value of the unbalanced energy norm for the four different balancing methods. Notice the similarity between the values obtained using the unbalanced fields from the OPV balance routine with fixed contours and those obtained using the first order δ–γ balance. We shall see that these two methods decompose the flow into almost indistinguishable balanced and unbalanced components.

4.2 Balanced dynamics

Across the four cases, the OPV balance procedure proves most effective, at least for appropriately chosen $\Delta\tau_D$. That is, the greatest fraction of the flow is attributed to balance, i.e. to motions induced by the PV field alone. The first-order δ–γ balance attributes significantly less of the flow to balance, and it fails to converge for case 3. This lack of convergence also occurs in the related f-plane context (Mohebalhojeh & Dritschel, 2001), where it was shown that the

Case	$\tilde{h}_\mathrm{i}/\tilde{h}$	δ_i/δ	ζ_i/ζ	γ_i/γ	v_i/v	u_i/u
(1)	0.0009	0.0167	0.0001	0.0283	0.0002	0.0002
(2)	0.0065	0.3811	0.0013	0.0229	0.0005	0.0006
(3)	0.0814	0.6041	0.0340	0.7538	0.0568	0.0458
(4)	0.0361	0.2837	0.0145	0.1778	0.0127	0.0099

Table 4.2: The magnitudes of the residual unbalanced fields over the first 10 days of the simulation using the OPV balancing routines. The values correspond to ratios of time-mean r.m.s. values.

order of the balance which minimises measures of imbalance is inversely related to the Froude number. In other words, higher-order balance relations such as $\delta_{tt} = \gamma_{tt} = 0$ might perform better than first-order balance for small Froude numbers as in case 1, but may not for moderate Froude numbers. The OPV balance does not appear to be as limited. Indeed, the OPV balance with fixed PV contours appears to be virtually equivalent to the the first-order δ–γ balance — without the problems of convergence. This is significant.

The magnitude of the residual flow (that which cannot be attributed to the PV) is listed in Table 4.2 for the OPV balance fields. The divergence (δ) and acceleration divergence (γ) fields have the greatest level of imbalance, especially in case 3. On the other hand, the depth and velocity fields are dominantly balanced. *Only a very small fraction of these fields is evolving independently of the PV.*

Using the definition of the imbalanced energy norm in equation 4.1.2, we can examine the time evolution of the imbalance (figure 4.2). Generally we see that $||E_i||$ increases dramatically with Fr (compare cases 1 and 3 or cases 2 and 4) and less so with Ro (compare cases 1 and 2 or cases 3 and 4), independently of which balance routine is used to produce the results. Again, however, the full OPV balancing routine performs best in all cases, at all times. The OPV routine also

reveals some interesting trends in the imbalance which are not noticeable in the other results. For example, in case 2 the imbalance decays over the first 4 days before levelling out, suggestive of an adjustment process taking place. Conversely, in case 3, the imbalance rapidly increases at the beginning of the simulation, but again levels out after around 4 days and remains relatively small. Remember that the initial conditions are merely random perturbations to the background planetary rotation – we have not prescribed any initial balance. Although the flows have been smoothly ramped up in order to reduce the imbalance it appears that the different timescales involved in each case has an effect on the amount of imbalance remaining after the end of the ramp period.

We examine next the spatial structure of the balanced and unbalanced fields defined by OPV balance. The balanced fields generally differ little from the full fields shown in the previous section with the exception of δ and γ in case 3. This is illustrated in Figure 4.3 for the divergence field, its balanced and unbalanced components (left, middle and right columns) for all four cases at time $t = 5$. Only in case 3 do differences in δ and δ_{b} become significant. The unbalanced field δ_{i} reveals wave-train structures highly suggestive of gravity-wave propagation, clearly visible in movies of the flow evolution. These waves are strong enough to show up also in the height field $\tilde{h}$ (Figure 4.4).

The results obtained from the OPV balance routine are now compared with those obtained using the first two members of the δ–γ hierarchy. Please note the contour intervals given above each frame in figures 4.5 - 4.7. They have been chosen, as far as possible, so that figures can be easily compared, but the density of contours sometimes makes this difficult. We focus on the results at $t = 5$ but the general picture remains the same throughout the evolution of the flow. The balanced fields differ little between the three methods, apart from the δ and γ fields which are considered to be entirely unbalanced by the zeroth order δ–γ balance ($\delta = \gamma = 0$). The differences become apparent when the

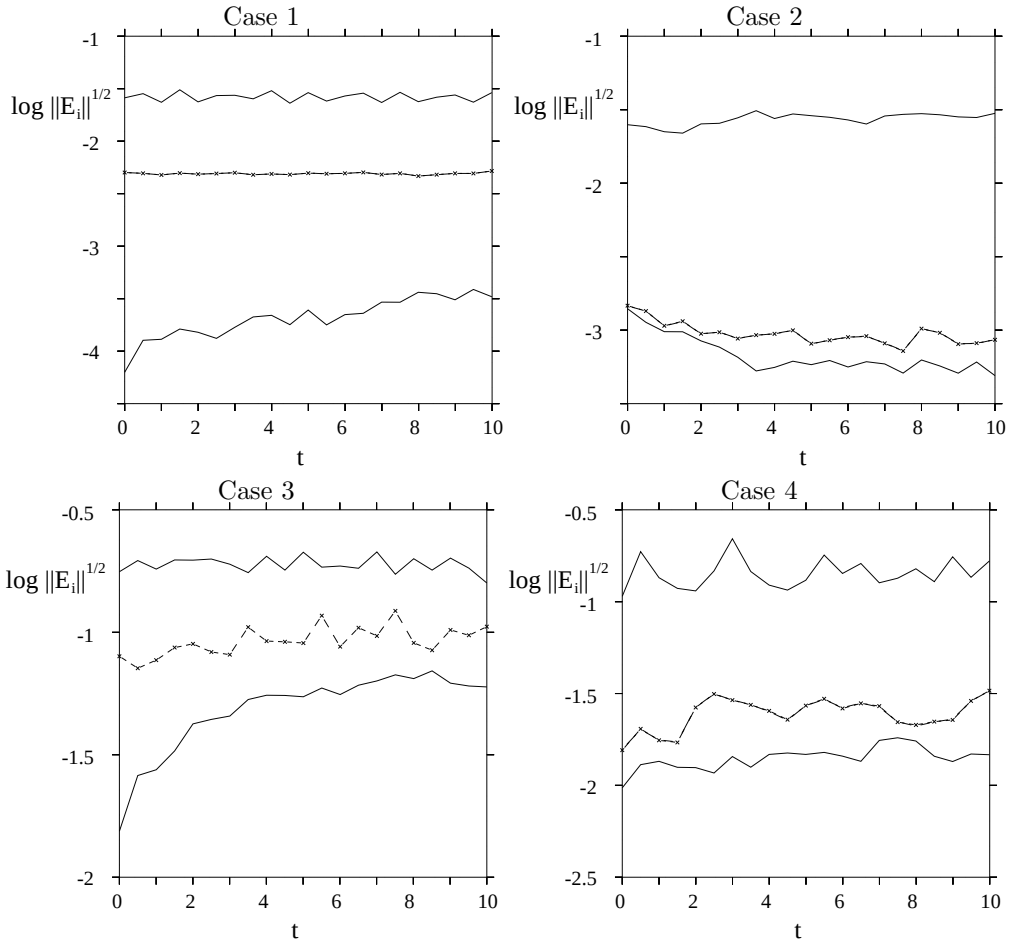


Figure 4.2: Time evolution of the imbalanced energy norm calculated using the four different balance routines: OPV (bold line), OPV with fixed contours (dash-cross-dash line), first order δ–γ balance (dashed line) and zeroth order δ–γ balance (solid line). Note that for case 3 the first order δ–γ balance (dashed line) is absent as this method did not converge for this case.

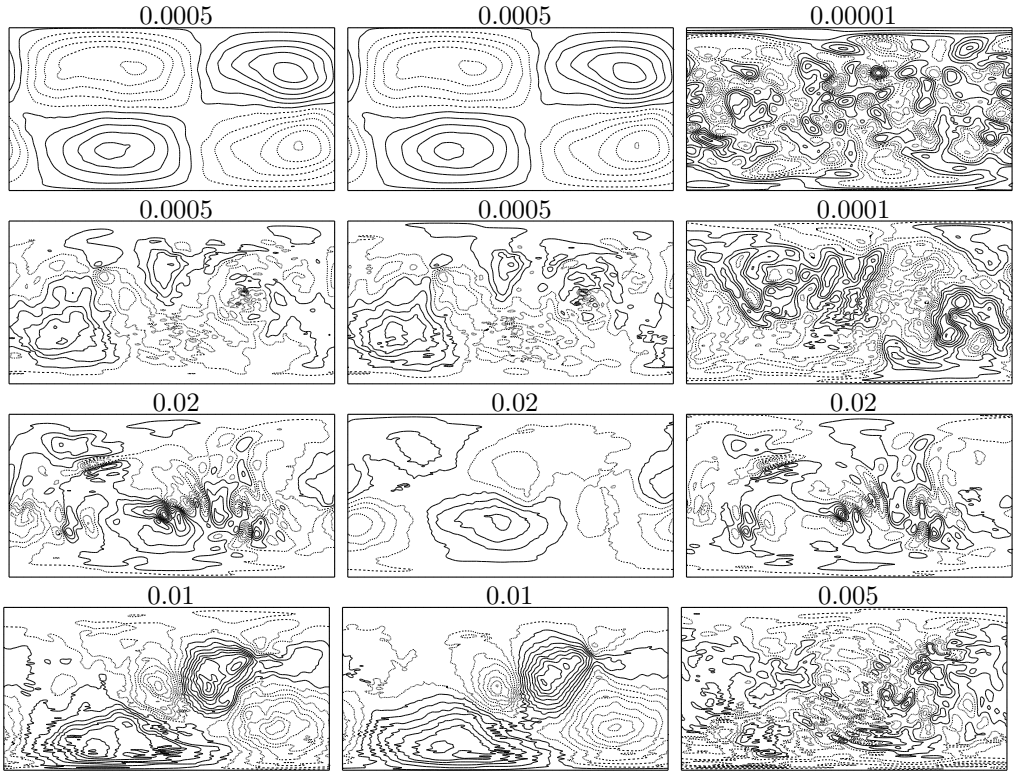


Figure 4.3: Full, OPV balanced and unbalanced components (left, middle and right columns) of $\delta/(2\Omega)$ at time $t = 5$ for cases 1–4 (top to bottom rows). The contour interval is shown above each image.

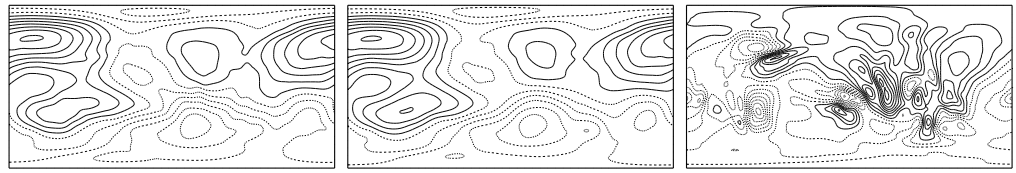


Figure 4.4: Full, OPV balanced and unbalanced components (left, middle and right columns) of $\tilde{h}$ at time $t = 5$ for case 3. The contour interval is 0.05 for $\tilde{h}$ and $\tilde{h}_{\mathsf{b}}$, and 0.005 for $\tilde{h}_{\mathsf{i}}$.

unbalanced fields are compared (figures 4.5 - 4.7). In general, both the zeroth and first order δ–γ balance conditions result in unbalanced fields that are larger scale and smoother than the corresponding OPV unbalanced fields. The large scale structure may in part be attributed to the equatorial wave which is obscured by the δ–γ based balance conditions. This will be clarified below when we analyse frequency spectra of the different fields. Some of the other features, particularly in the zeroth order δ–γ unbalanced fields, can be traced back to the structure of the full field, for example, compare the unbalanced depth field given by zeroth order δ–γ balance with the corresponding full depth field in figure 3.7. This is an indication that the δ–γ balance conditions give unbalanced fields that in fact still contain a significant amount of balanced motion.

Case 3 provides us with a good example of how the different balancing routines cope with gravity waves. We have seen in figure 3.8 that a gravity wave dominates the divergence field at $t = 5$. This wave also shows up clearly in all three definitions of the imbalanced divergence field. A stricter test is how well the wave can be seen in the other unbalanced fields where it is not such a dominating feature. Here the OPV balance routine performs best as the wave can be clearly seen in both the depth and meridional velocity fields (figures 4.5 and 4.7). The wave can also be seen in the depth and meridional velocity fields defined by first order δ–γ balance but the presence of high magnitude large scale structure obscures it. Zeroth order δ–γ balance considers the entire divergence field to be unbalanced so does a poor job of separating the wave at all.

One way of examining the large scale structure remaining in the unbalanced components it to look at how the magnitude of the imbalance depends on latitude. There are two issues here. Firstly, the accuracy of the balancing routine might be dependent on latitude. For example, we expect the balancing routines based on the δ, γ variables to perform better away from the equator since the underlying geostrophic balance on which they are based potentially breaks down as $f \to 0$.

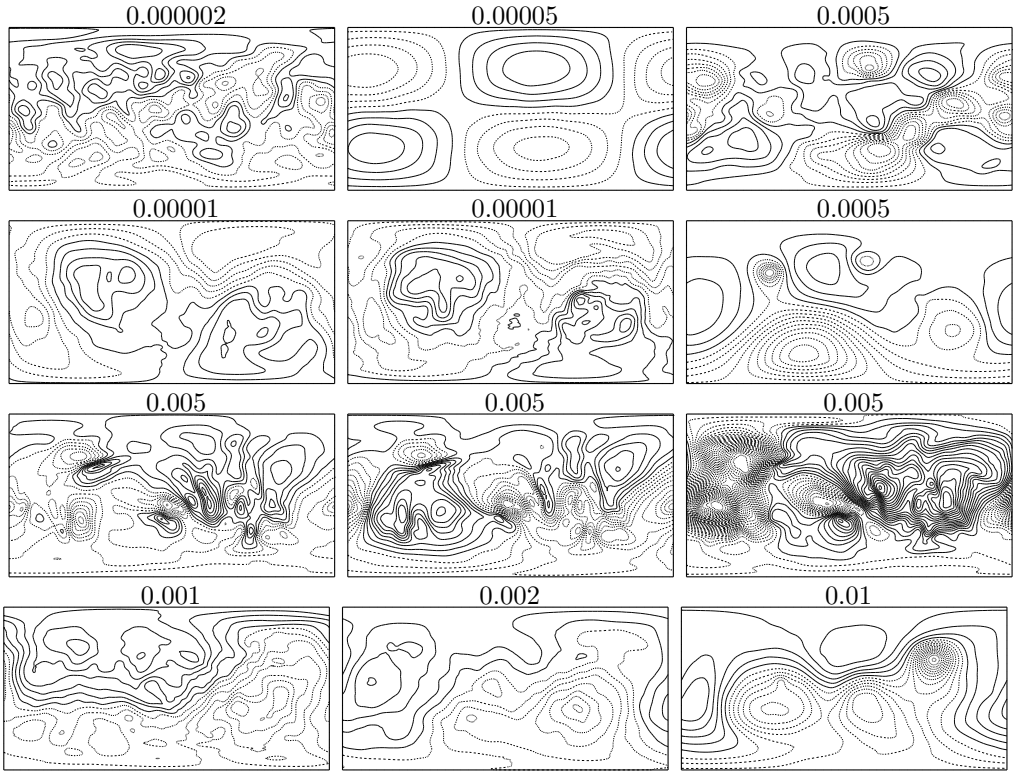


Figure 4.5: Fields of $\tilde{h}_{\mathrm{i}}$ at t=5 for the four cases. The result from the OPV balance routine is shown on the left, first order δ–γ balance in the middle (except for case 3 where the OPV balance fixed contour routine results are used instead of the first order δ–γ balance results) and zeroth order δ–γ balance on the right.

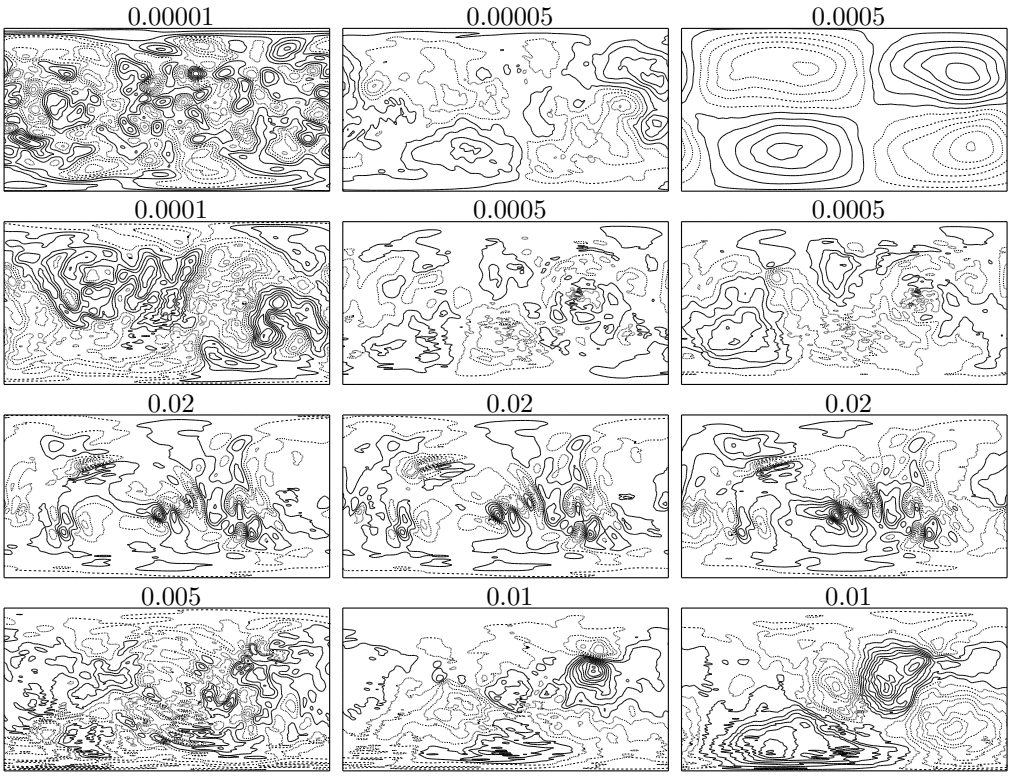


Figure 4.6: Fields of δ_i at t=5 for the four cases. The result from the OPV balance routine is shown on the left, first order δ–γ balance in the middle (except for case 3 where the OPV balance fixed contour routine results are used instead of the first order δ–γ balance results) and zeroth order δ–γ balance on the right. Note that the right figure is identical to the full δ field as the zeroth order δ–γ balance considers the entire δ field to be unbalanced.

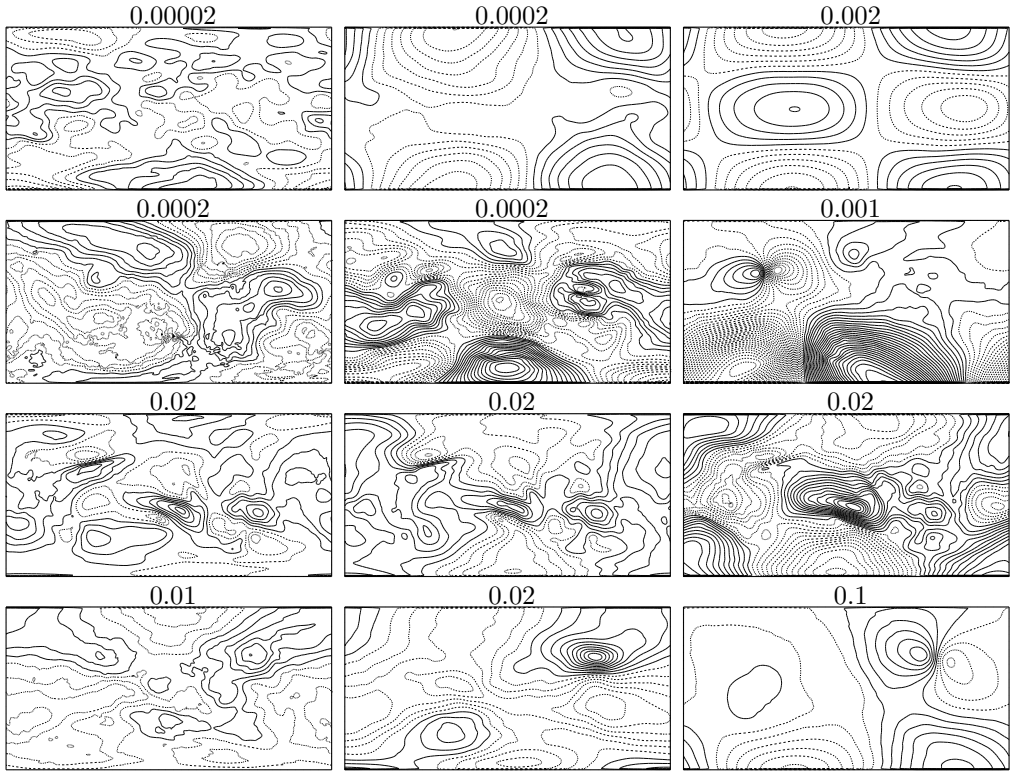


Figure 4.7: Fields of meridional velocity v_{i} at t=5 for the four cases. The result from the OPV balance routine is shown on the left, first order δ–γ balance in the middle (except for case 3 where the OPV balance fixed contour routine results are used instead of the first order δ–γ balance results) and zeroth order δ–γ balance on the right.

Secondly, inertia-gravity waves may be generated and confined in certain regions, e.g. the inertia-gravity wave generated to the south of the equator in case 3. In figures 4.8-4.10 we plot the r.m.s. value of the unbalanced depth, divergence and acceleration divergence fields, averaged over longitude and time, versus latitude. In general, the OPV unbalanced fields show the least dependence on latitude. For cases 3 and 4 the OPV unbalanced fields show increased magnitude to the south of the equator, corresponding to the actual unbalanced motion in this region (see figure 4.6). The OPV unbalanced fields are generally weaker than those produced by the other balancing routines, although in some cases ($\tilde{h}_\mathrm{i}$ in case 2 and γ_i in cases 3 and 4) the difference is slight. Case 1 is an interesting example of the large scale structure remaining in the unbalanced depth fields produced by the balance routines based on the δ and γ variables. The zeroth order routine sees a global scale wave in the divergence field whereas the first order routine sees this wave in the depth field and, to a lesser extent, in the γ field. In some cases it is possible to say that the zeroth and first order δ–γ balancing routines perform worse at the equator than in the midlatitudes (see the figures for case 4 in particular and the zeroth order δ–γ unbalanced depth field for case 1) but in general the difference is not great. This is significant as these routines, based on a variant of geostrophic balance, were not expected to perform so well across the whole sphere. Further discussion may be found in Smith and Dritschel [2006].

4.3 Frequency spectra

We now consider the frequency spectra of the balanced and unbalanced components of the flow for the four cases. These have been calculated from time-series of the data, saved at every grid point at time intervals of 0.001 days, for the first 10 days of each simulation. The 128^2 spectra are averaged by integrating over the area of the sphere.

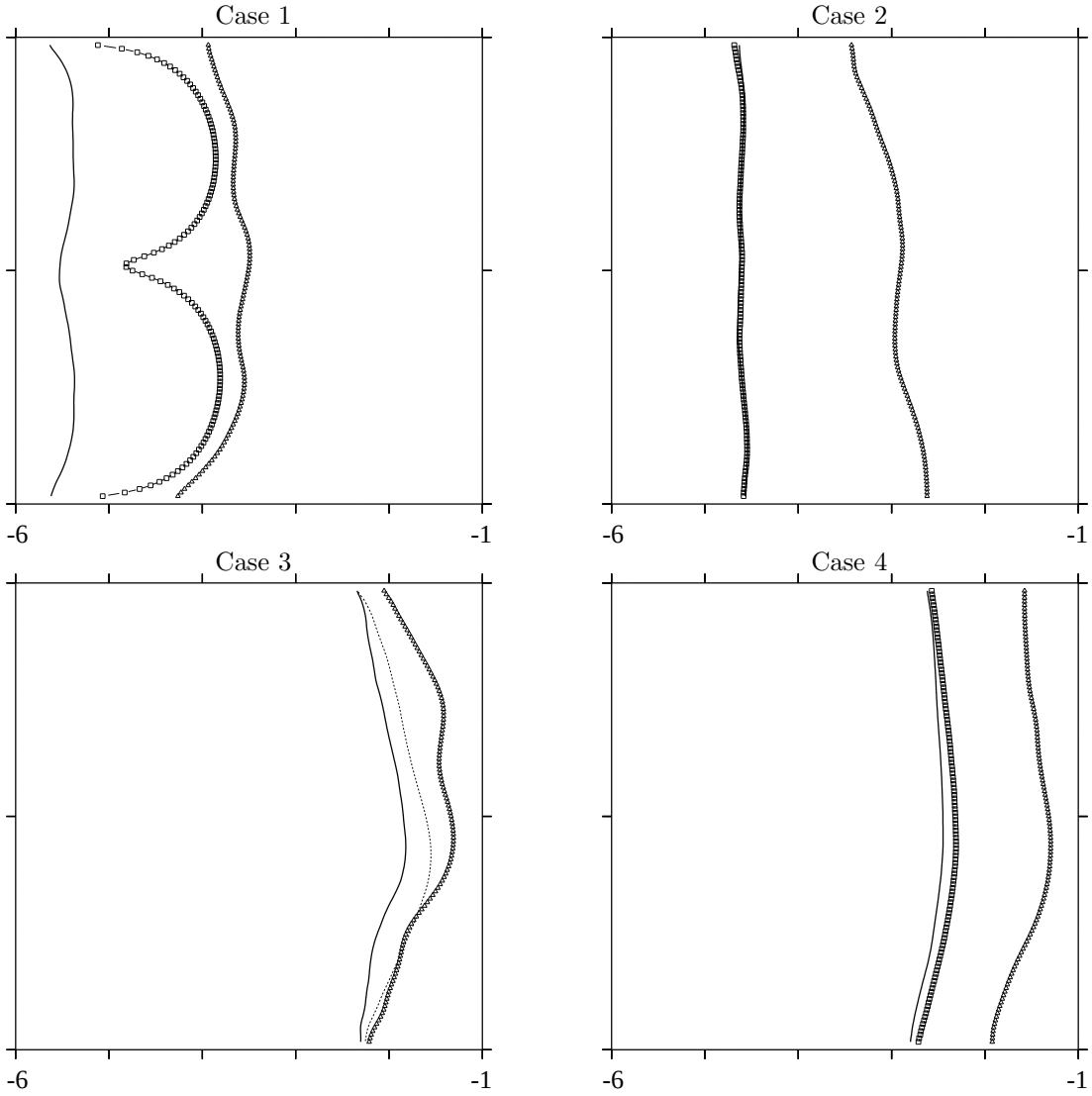


Figure 4.8: The $\log_{10}$ of the r.m.s. value of the unbalanced depth field, averaged over longitude and time, (x axis), plotted versus the latitude (y axis). The results are for OPV balance (bold line), OPV balance with fixed contours (dashed line), first order δ–γ balance (squares) and zeroth order δ–γ balance (triangles). The dashed line (OPV balance with fixed contours) and the squares (first order δ–γ balance) lie on top of each other.

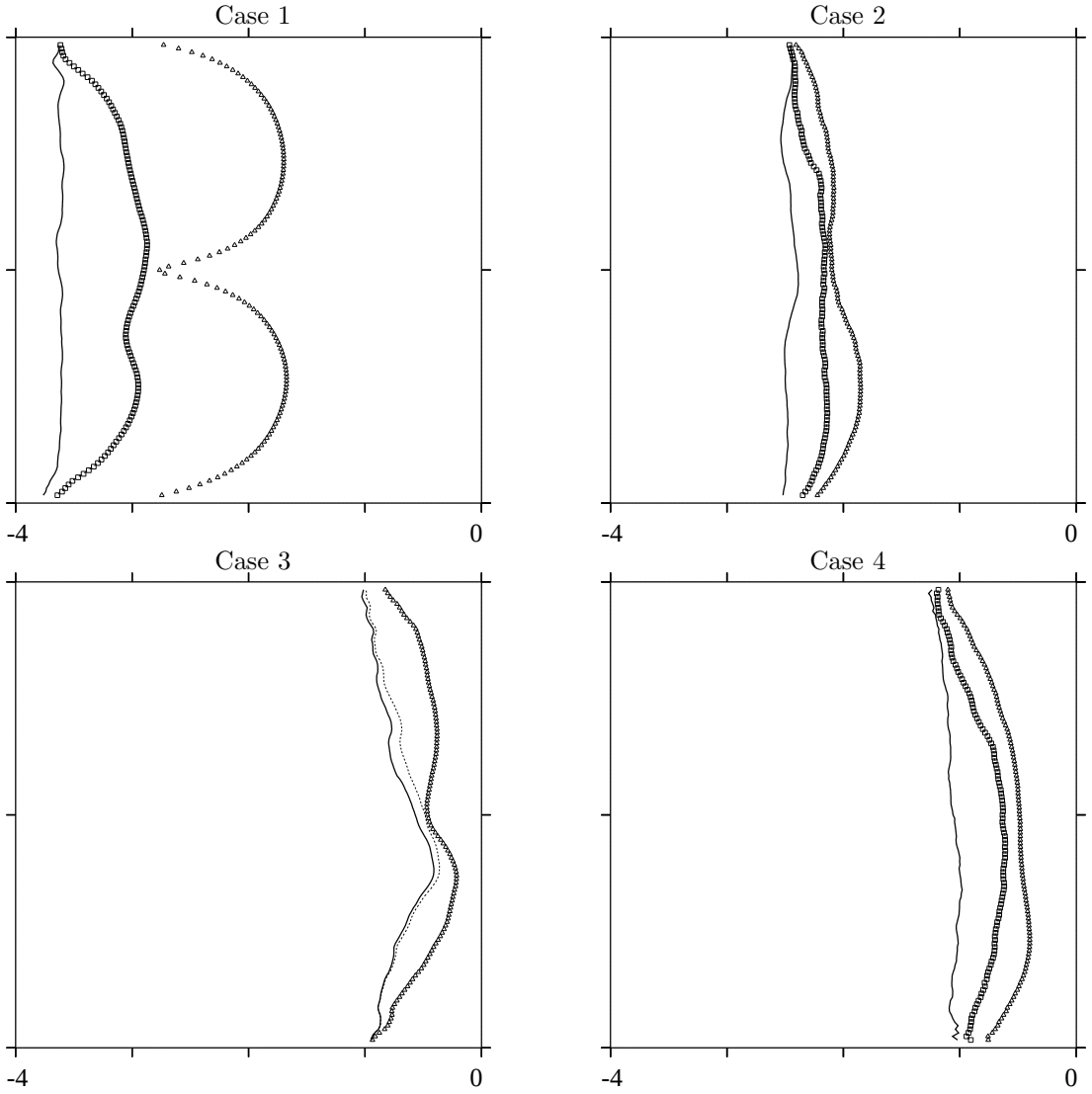


Figure 4.9: The $\log_{10}$ of the r.m.s. value of the unbalanced divergence field, averaged over longitude and time, (x axis), plotted versus the latitude (y axis). The results are for OPV balance (bold line), OPV balance with fixed contours (dashed line), first order δ–γ balance (squares) and zeroth order δ–γ balance (triangles). The dashed line (OPV balance with fixed contours) and the squares (first order δ–γ balance) lie on top of each other.

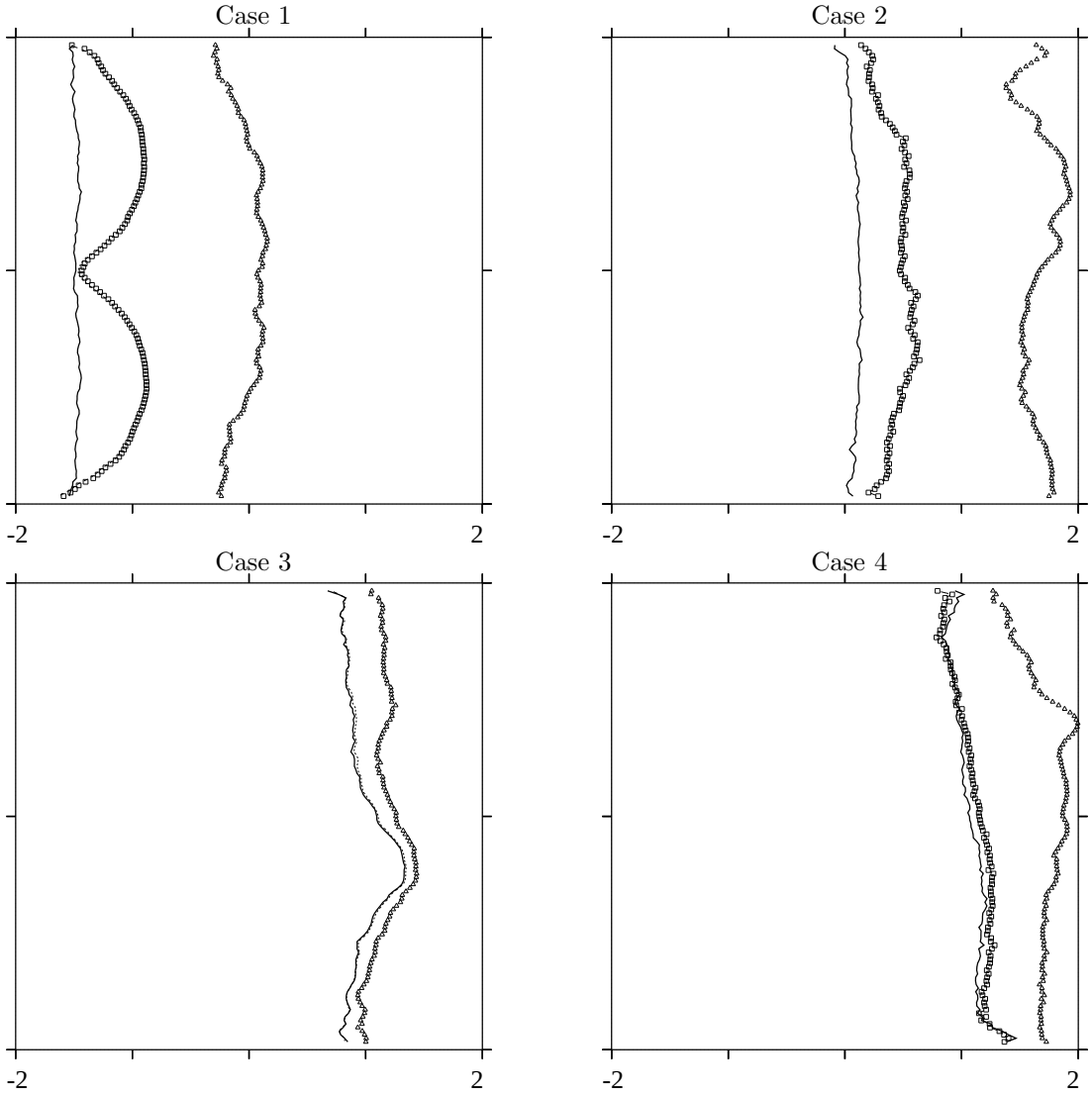


Figure 4.10: The $\log_{10}$ of the r.m.s. value of the unbalanced acceleration divergence field, averaged over longitude and time, (x axis), plotted versus the latitude (y axis). The results are for OPV balance (bold line), OPV balance with fixed contours (dashed line), first order δ–γ balance (squares) and zeroth order δ–γ balance (triangles). The dashed line (OPV balance with fixed contours) and the squares (first order δ–γ balance) lie on top of each other.

First we examine case 1. Figure 4.11 shows the frequency spectra of the depth field and its balanced and unbalanced components calculated using the four balancing routines. In this case the flow remains close to balance and the depth field is the most balanced field so we would expect the differences between the balancing routines to be minimal. All the routines produce balanced spectra that are indistinguishable by eye from the spectra of the full field. However, the differences between the unbalanced spectra are significant. The full OPV balance scheme achieves the greatest separation between the balanced and unbalanced spectra at all frequencies but with the most improvement occurring at low frequencies. It is also the only balance scheme that nearly removes all trace of the equatorial wave (at $\omega = 1$) from the unbalanced spectra. Again the results produced by OPV balance with fixed PV contours are essentially identical to those produced by first order δ–γ balance. Note the relatively large erroneous imbalance at low frequencies suggested by the zeroth order δ–γ balance.

Having established that OPV balance achieves the greatest separation between the balanced and unbalanced frequency spectra, we go on to examine the OPV balanced and unbalanced divergence spectra for all 4 cases (see figure 4.12). Case 3, the case least representative of geophysical flows, is the case that best demonstrates the classic features that we have come to expect from the traditional picture of slow-vortical and fast-gravity wave separation. For $\log_{10}|\omega| < 0.3$ the balanced motion dominates. At $\log_{10}|\omega| = 0.3$ the balanced and imbalanced spectra cross over and the imbalance dominates at the high frequency end of the spectrum. The balanced and imbalanced spectra also cross over in case 4, this time around $\log_{10}|\omega| = 0.6$. However, in this case, the magnitude of the unbalanced spectra does not vary significantly with frequency, i.e. the imbalance is not confined to high frequencies. The unbalanced spectrum in case 2 is also relatively flat but with peak around $\log_{10}|\omega| = -0.4$, i.e. in the low (sub inertial) frequency range. In case 1 the balanced and unbalanced spectra do not cross, or

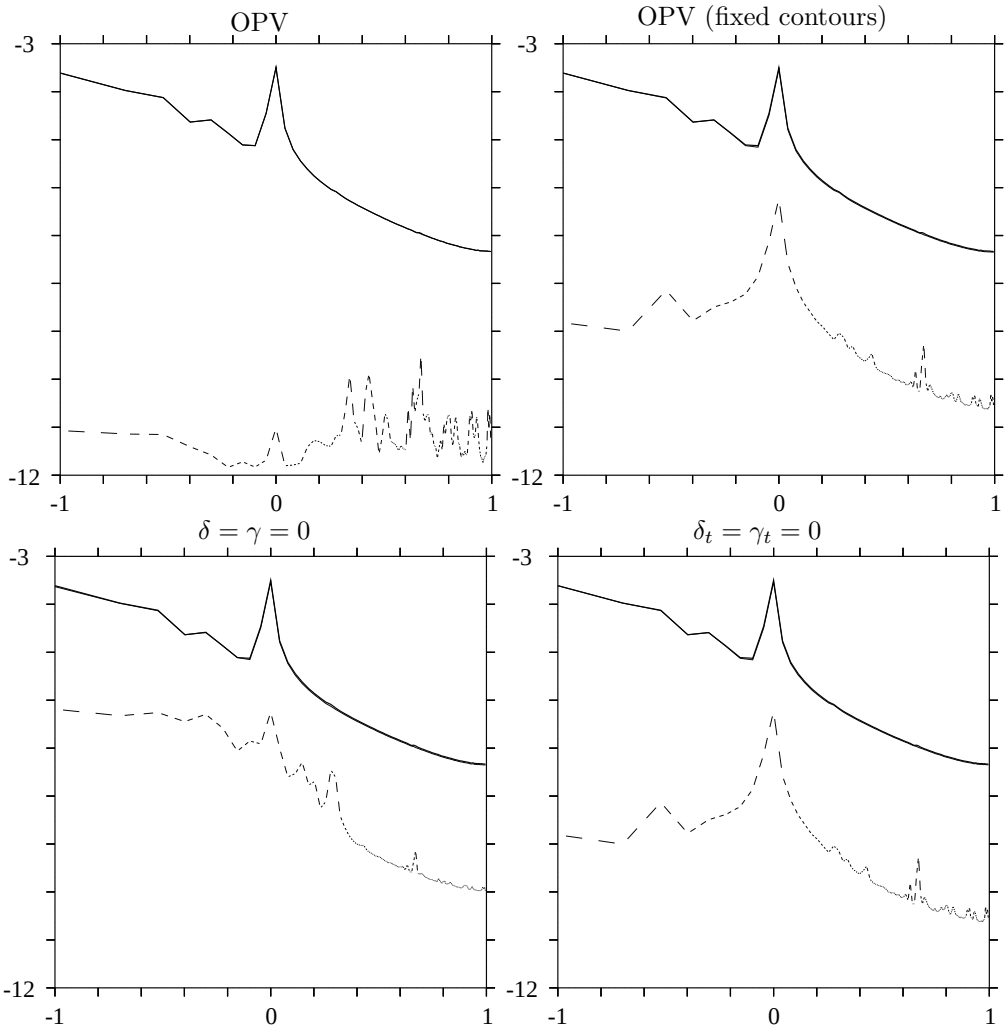


Figure 4.11: Frequency spectra $\log_{10}|\hat{\xi}|^2$ for $\xi = \tilde{h}$ (bold line), $\xi = \tilde{h}_{\mathsf{b}}$ (thin line) and $\xi = \tilde{h}_{\mathsf{i}}$ (dashed line) versus $\log_{10}(\omega)$ for case 1. The spectra of $\tilde{h}$ and $\tilde{h}_{\mathsf{b}}$ (bold and thin lines) are indistinguishable.

even come close to each other. The balanced spectrum dominates at all scales. These results show that there is no clear frequency separation between balanced and unbalanced motion. *Separating a flow into fast and slow frequency motion is not equivalent to decomposing vortical and gravity wave motion.* Thinking in terms of wave–vortex decomposition is more accurate and informative.

4.4 Parameter space results

So far we have examined four of our simulations in detail. We now examine the entire Fr-Ro parameter space. Given the large number of simulations it is useful to define a single number that will tell us something about each flow that we can then use to produce a contour plot over the parameter space. For example, we can measure the average magnitude of each field over the sphere for the duration of the simulation. More usefully, we can measure the average magnitude of the unbalanced fields and compare this to the average magnitude of the full fields. This will give us a measure of how unbalanced the flows are and how this depends on the Froude and Rossby numbers of the flow.

Since the simulations do not fall on a regular grid in the Fr-Ro parameter space, we have to interpolate the data. Firstly we generate a regular grid x_i, y_j then for each grid point we loop over all data points f_k and compute

$$\frac{\Sigma_k f_k w_k}{\Sigma_k w_k}, \tag{4.3}$$

where w_k a weight function defined as

$$w_k = \exp\Big(-\frac{r^2}{2R^2}\Big), \tag{4.4}$$

where $r^2 = (x_i - x_k)^2 + (y_j - y_k)^2$ is just the square of the radial distance between

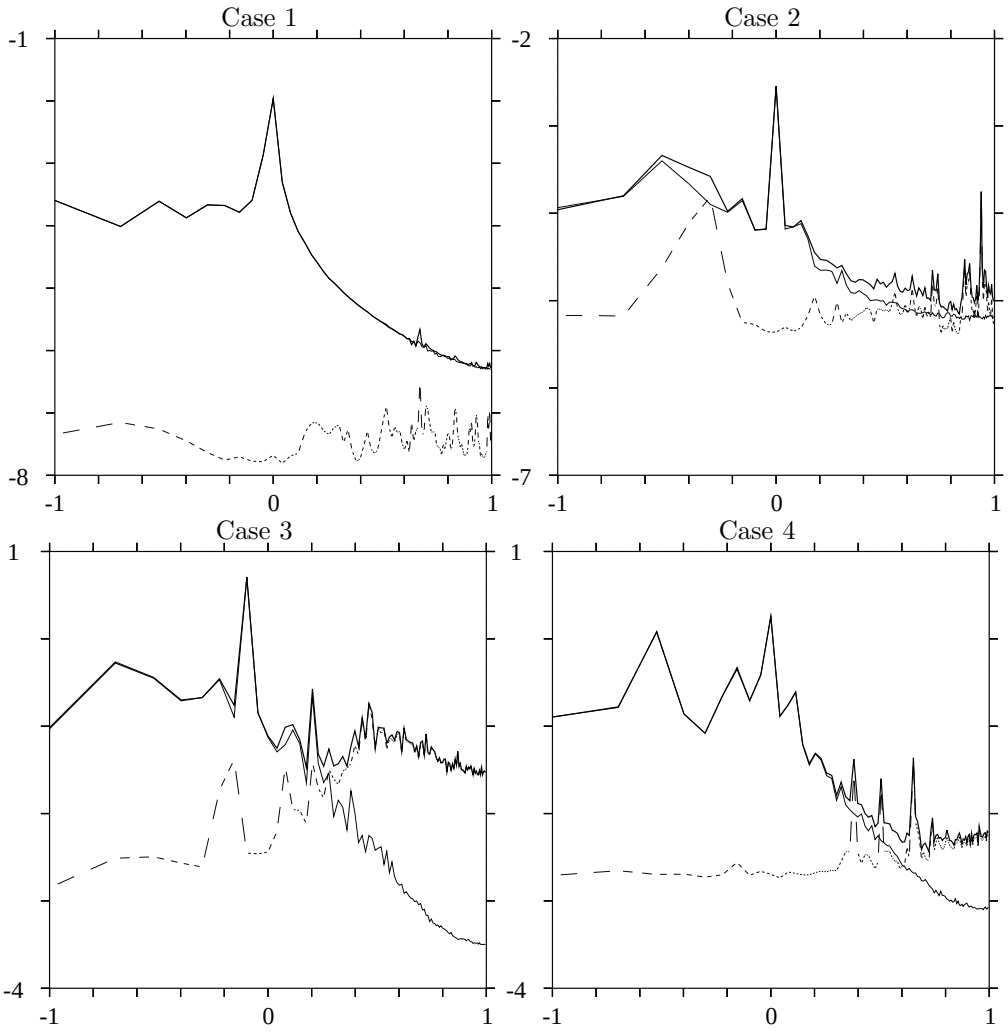


Figure 4.12: Frequency spectra $\log_{10}|\hat{\xi}|^2$ versus $\log_{10}(\omega)$ for $\xi = \delta$ (bold line), $\xi = \delta_{\mathsf{b}}$ (thin line) and $\xi = \delta_{\mathsf{i}}$ (dashed line) calculated using OPV balance.

the data point location (x_k, y_k) and the grid point location (x_i, y_j) and R is a predefined 'radius of influence' that controls how much data points contribute to the interpolated value at (x_i, y_j). We run this interpolation algorithm for several values of R and assess which gives the most accurate representation of the original data by bilinearly interpolating the data back to the original data point locations (x_k, y_k) and computing the rms average of the difference between the original data and the interpolated data. The figures shown in this section are those judged by this method to be the best representation of the original data, i.e. those for which the rms average difference between the original and interpolated data is minimal.

Figure 4.13 shows $\bar{q}$, where the overbar denotes the rms average, for $q = \tilde{h}, \delta, \zeta$ and γ over the parameter space. These figures show the balanced component of these fields but are indistinguishable by eye from those showing the full fields. We see that the average depth perturbation and the average divergence both increase with Froude number. This is unsurprising since the greater the Froude number, the more deformed is the fluid surface. The two vorticity based variables both increase with Rossby number; again unsurprising as the vorticity scales with Rossby number.

In order to quantify the extent to which flows remain balanced we compute the ratio the rms average of the unbalanced field to the r.m.s. average of the full field i.e. $\bar{q}_{\mathrm{i}}/\bar{q}$. Again we produce contour plots over the Fr-Ro parameter space (see figures 4.14-4.16). According to the OPV routines the imbalance in the depth field increases in general with Froude number although there is a region of parameter space near the bottom boundary (where Ro $\sim$ 2Fr) where flows are less balanced. The imbalance in the divergence field follows a similar pattern of increase with Froude number, although the magnitude of the imbalance is much greater. The increase in imbalance with Froude number is less clear, but still present, in the vorticity field plots (figure 4.15). However, the figure for the zeroth order δ–γ balance (figure 4.16) suggests that the unbalance in the vorticity field is more

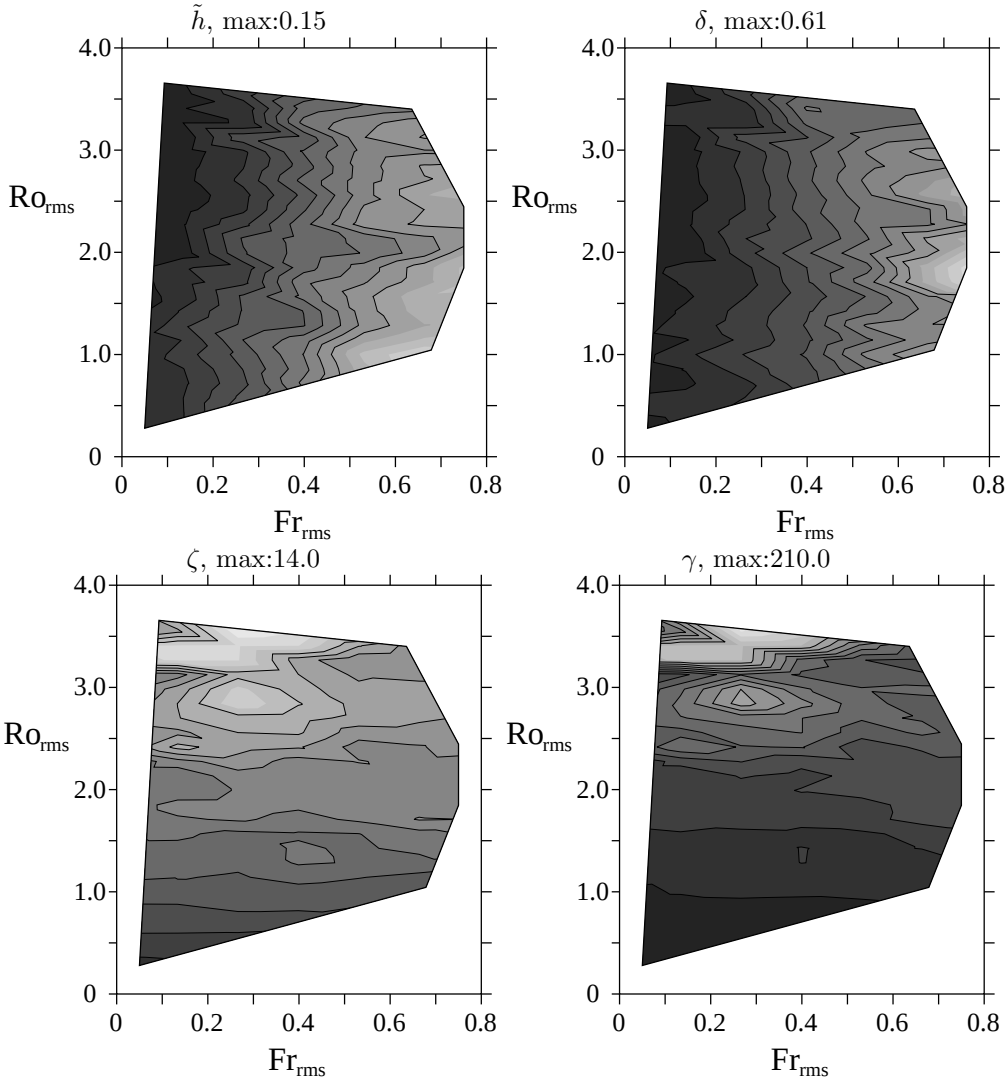


Figure 4.13: Contour plots over the Fr-Ro parameter space of the magnitude of the OPV balanced depth, divergence, vorticity and gamma fields, all plotted with 15 contour intervals. The maximum value plotted is shown above each figure. White indicates large values, black indicates small values.

dependent on Ro. This illustrates how different definitions of balance can lead to totally different conclusions and in particular how zeroth order δ–γ balance — an improvement on geostrophic balance — can be seriously misleading.

4.5 Impact of imbalance

In the previous sections we have compared the balanced component of the full flow to the flow computed using a balanced model. This gives us an indication of how well the balanced models capture the features of the full flow. We can also see, for example in figure 4.2, how the importance (i.e. magnitude) of the imbalance changes with time. In this section we examine what impact the unbalanced motion has on the flow. From figure 4.2 we see that it is possible for the flow to adjust towards a more balanced state (see case 2) or conversely for the flow to generate imbalance which then grows (see cases 1 and 3). We now ask whether the unbalanced motion can modify the balanced flow. In the setup used so far it is possible for the imbalance to accumulate over time. In order to better quantify the impact of the unbalanced motion on the flow we take each case and break it up into segments, compute the OPV balanced fields at the start of each segment and use these to re-initialise the flow. Then at the end of the segment we run the OPV balance routine again and compare the resulting fields with those computed using the OPV balance routine on the data from the initial simulation. To recap, we take the model state at time τ, $S(\tau)$, and compute the balanced component $S_b(\tau)$ using the OPV balancing routine. We then use this balanced component as the initial condition and run the model to $t = \tau + d\tau$ to obtain $S'(\tau + d\tau)$, where the prime indicates a new model state different to the original state $S(\tau + d\tau)$ at this time. Now we use the OPV balancing routine to compute the balanced component of the new flow at $\tau + d\tau$, $S'_b(\tau + d\tau)$, and compare this to the OPV balanced component of the original flow, $S_b(\tau + d\tau)$ at

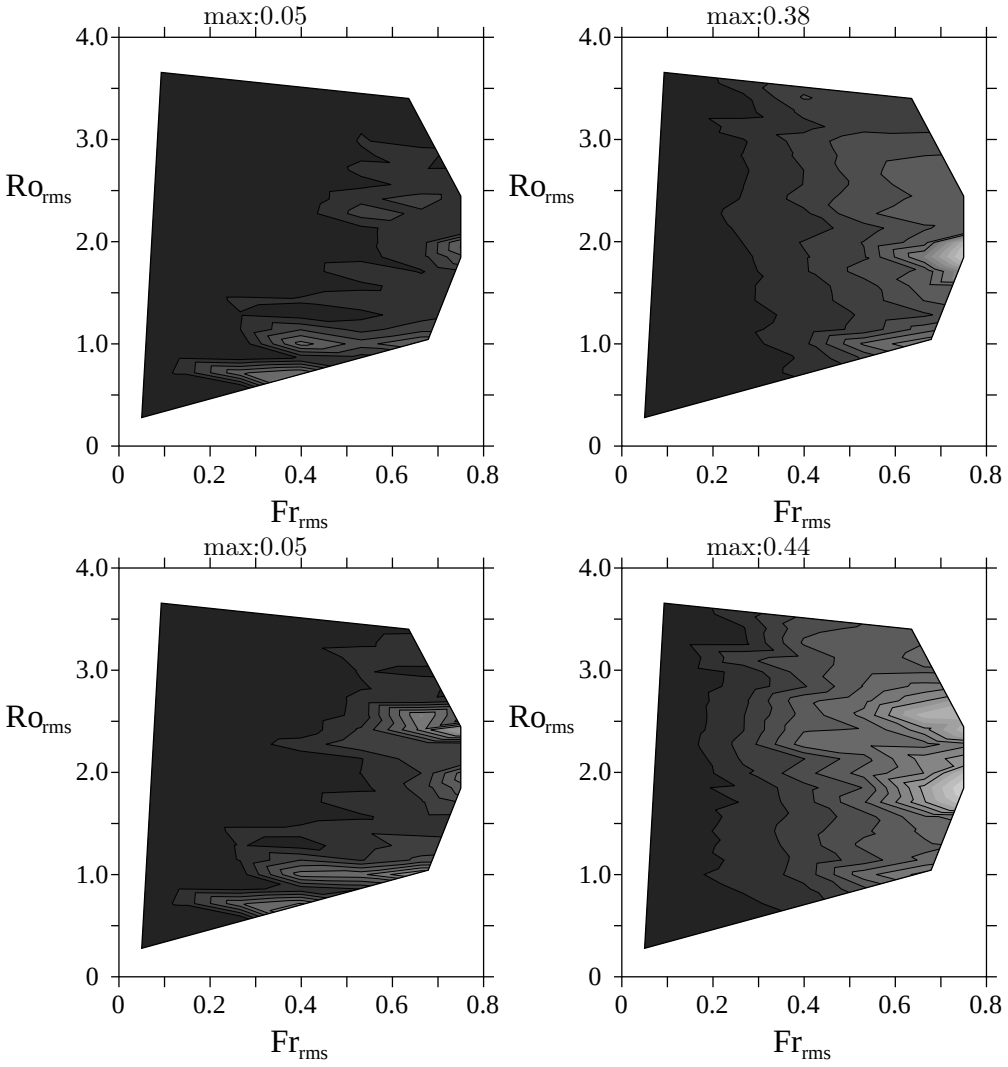


Figure 4.14: Contour plots over the Fr-Ro parameter space of $\bar{q}_i/\bar{q}$ where q is $\tilde{h}$ for the first column and δ for the second column. The first row uses the unbalanced fields produced by the OPV balancing routine, the second row those from the OPV routine with fixed contours. All are plotted with 15 contour intervals and the maximum value plotted is shown above each figure. White indicates large values, black indicates small values.

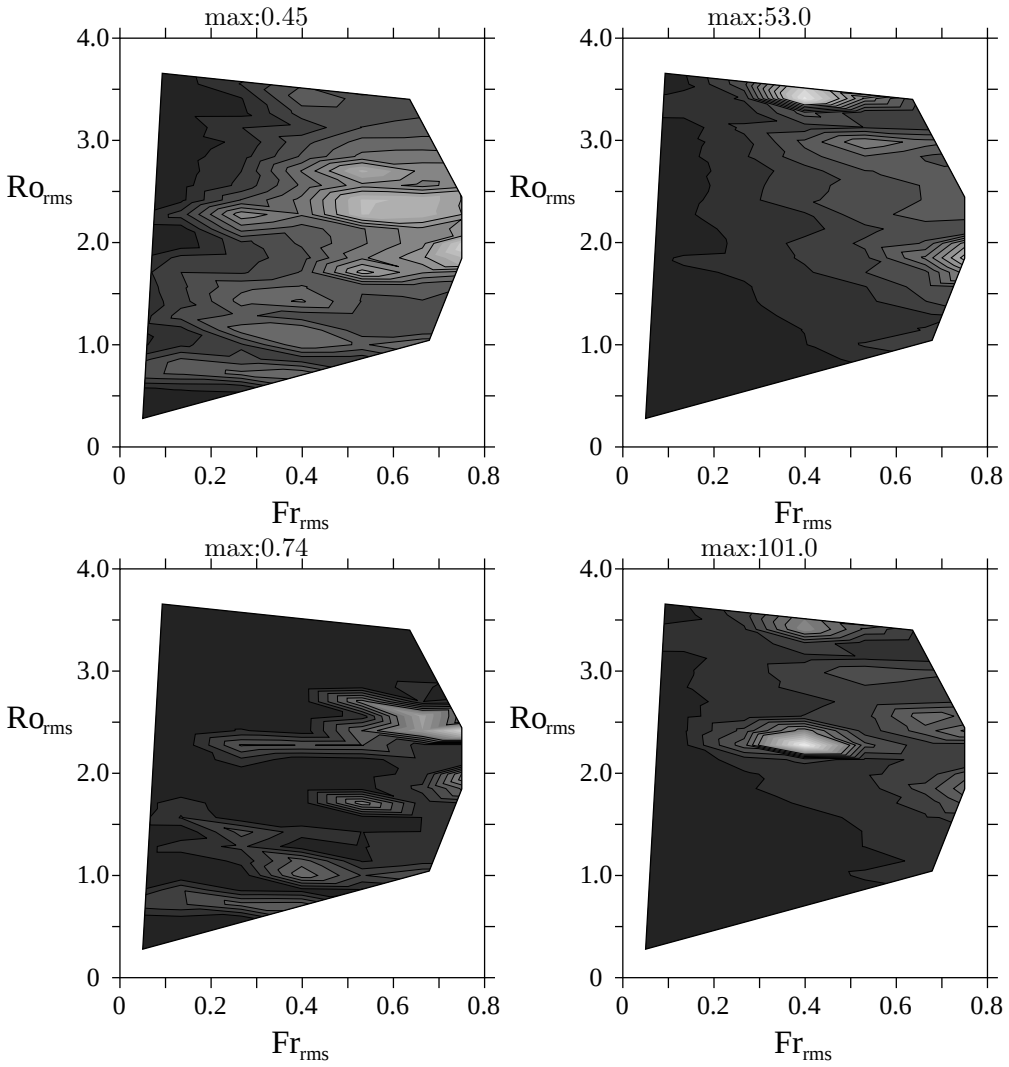


Figure 4.15: Contour plots over the Fr-Ro parameter space of $\bar{q}_i/\bar{q}$ where q is ζ for the first column and γ for the second column. The first row uses the unbalanced fields produced by the OPV balancing routine, the second row those from the OPV routine with fixed contours. All are plotted with 15 contour intervals and the maximum value plotted is shown above each figure. White indicates large values, black indicates small values.

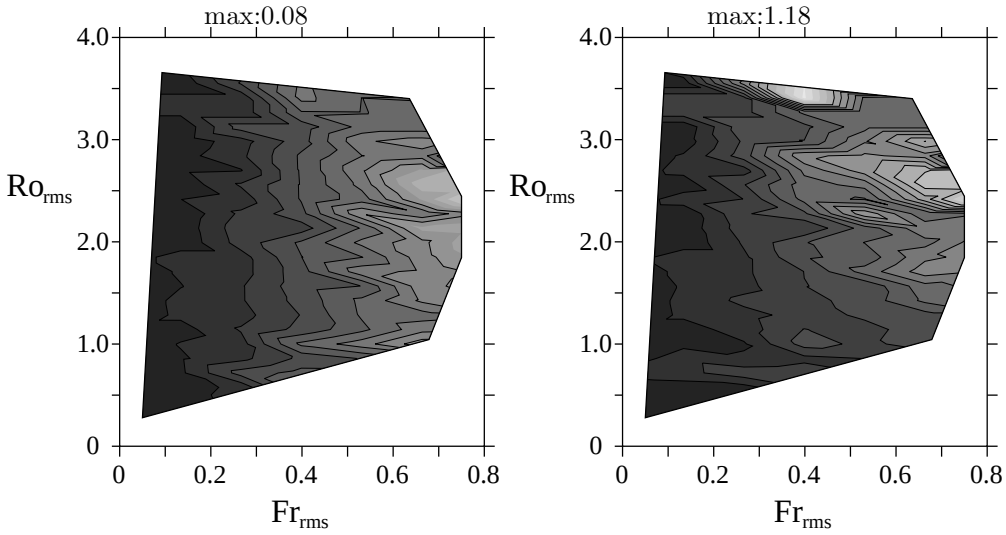


Figure 4.16: Contour plots over the Fr-Ro parameter space of $\bar{q}_{\rm i}/\bar{q}$ where q is $\tilde{h}$ for the first column and ζ for the second column, using fields from the zeroth order δ–γ balanced model. Both are plotted with 15 contour intervals and the maximum value plotted is shown above each figure. White indicates large values, black indicates small values.

the same time. Thus we are comparing the OPV balanced component of the flow produced by the original model run to the OPV balanced component of the flow computed from balanced initial conditions. If we repeat this at intervals of $d\tau$ we can compute the rate at which the unbalanced motion modifies the balanced motion. This process is outlined in the schematic diagram 4.17 and in the flow chart 4.18.

After plotting the r.m.s. difference between $\tilde{h}'_b(\tau)$ and $\tilde{h}_b(\tau)$ against τ we notice that the results are obscured by an oscillation with period t_{surg}, the surgical timescale. This is particularly pronounced in cases 1 and 2 (see figure 4.20 for an example). Figure 4.19 shows what happens to the results for case 1 if we switch off contour surgery. This removes the oscillation but unfortunately the simulation soon blows up due to the accumulation of small scales. A recent improvement to remove traces of surgery in the gridded PV field [Dritschel and Fontane] came too late to be implemented for the present purpose. However, it would be worthwhile repeating these simulations in the near future. For now, we notice from figure 4.19 that the results are very accurate if we ignore the times affected by surgery. In figure 4.21 we plot the r.m.s. difference between $\tilde{h}'_b(\tau)$ and $\tilde{h}_b(\tau)$ against τ for the times that are unaffected by surgery. In all cases the magnitude of this difference is very small, typically several thousandths of the magnitude of the full depth field, and remains small for all time. The impact of imbalance shows a significant decrease with time for case 2 and, to a lesser extent, for case 4. Cases 1 and 3 show the reverse with the impact of imbalance increasing slightly over time, although this tails off after $t = 4$ in case 3. This is the first time that this impact has been measured with the most accurate method available.

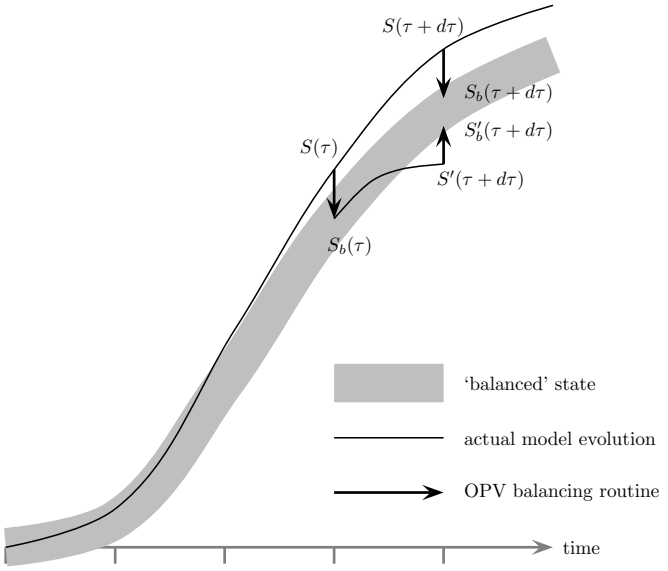


Figure 4.17: This is a schematic of the various model states and their balanced components. The solid line represents the state of the initial simulation. At $t = \tau$ we run the OPV balance routine (bold arrow) to push the model state onto the 'balanced manifold', represented by the thick grey line. The thickness of this line indicates the 'fuzziness' of this 'manifold'. This balanced state is used to initialise another simulation (shown as a solid line) which will not necessarily remain balanced. After a time interval $\Delta\tau$ we run the OPV balancing routine again. This will not give us the same balanced fields as we obtain from running the OPV routine on the original fields. Comparing the two different balanced states will give us a measure of how the unbalanced motion present in the original run has affected the balanced component of the flow.

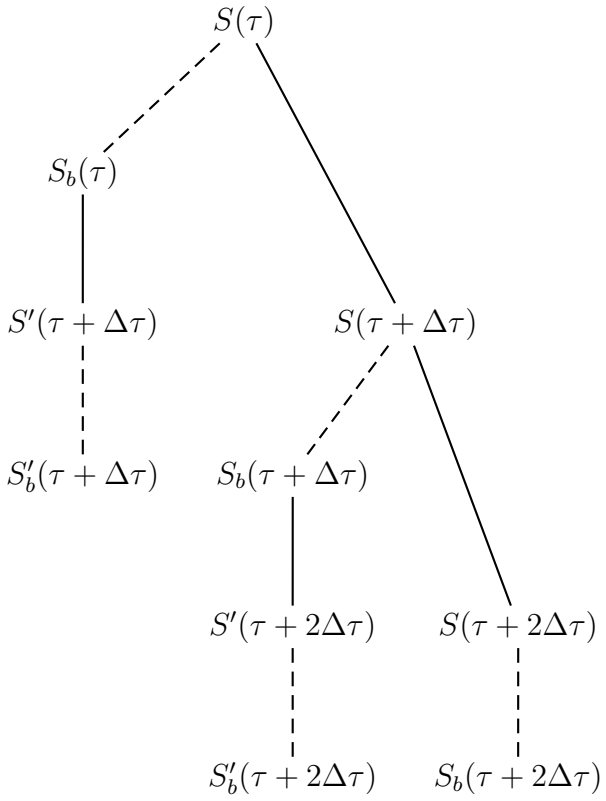


Figure 4.18: This flow diagram illustrates the process for computing the rate at which the unbalanced motion affects the balanced flow. The model state at time τ is represented by $S(\tau)$. $S'(\tau)$ indicates that this model state has been computed from OPV balanced initial conditions. A dashed line connecting $S(\tau)$ to $S_b(\tau)$ indicates that $S_b(\tau)$ is obtained from $S(\tau)$ by OPV balance. A solid line connecting $S(\tau)$ to $S(\tau + \Delta\tau)$ indicates that $S(\tau + \Delta\tau)$ is obtained from $S(\tau)$ by integrating the full CASL model. We start with the model state at $t = \tau$. The diagram shows how we then obtain $S'_b(\tau + \Delta\tau)$ and $S_b(\tau + \Delta\tau)$ which are the states we need to compare to find out how the unbalanced motion affects the balanced flow. The diagram also shows how $S'_b(\tau + 2\Delta\tau)$ and $S_b(\tau + 2\Delta\tau)$ are computed. For our analysis, we start from $\tau = 0$ and use $\Delta\tau = 0.05$.

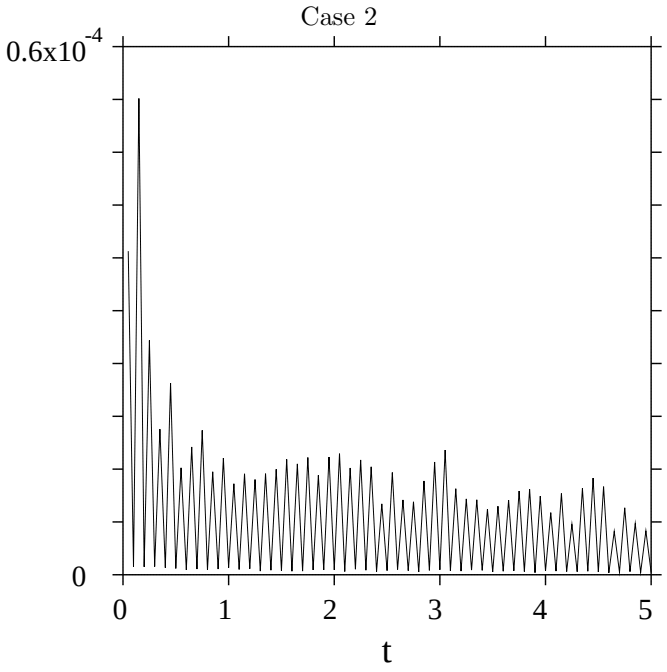


Figure 4.19: The r.m.s. difference between $\tilde{h}'_b(\tau)$ and $\tilde{h}_b(\tau)$ plotted versus τ for case 2.

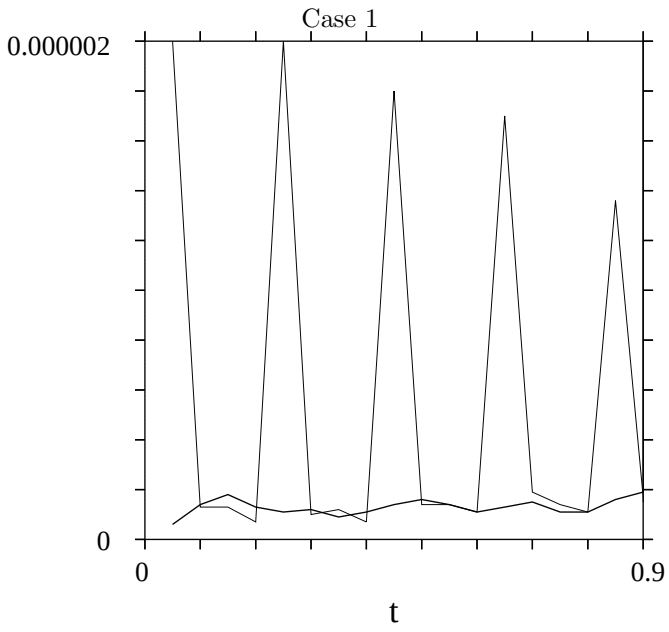


Figure 4.20: The r.m.s. difference between $\tilde{h}'_b(\tau)$ and $\tilde{h}_b(\tau)$ plotted versus τ for case 1. The solid line shows the original result; the bold line shows the result with contour surgery switched off.

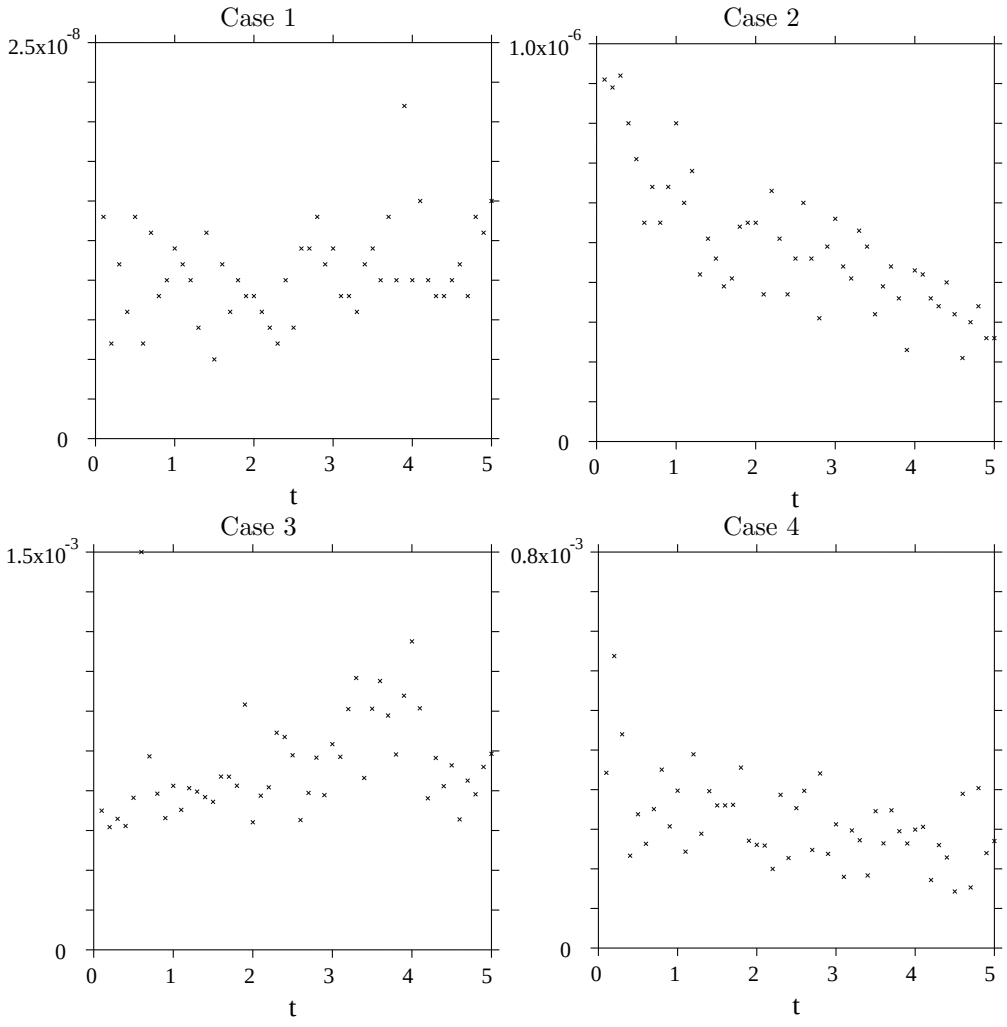


Figure 4.21: These figures show the r.m.s. difference between $\tilde{h}'_b(\tau)$ and $\tilde{h}_b(\tau)$ plotted versus τ for each of the four cases. There is one point in case 3 that goes off the scale - this point has a value of 0.015.

Chapter 5

Discussion and future work

This chapter contains a summary and brief discussion of the results presented in each previous chapter. We also present some suggestions for future extensions of this work.

In chapter 2 we outlined the mathematical and numerical setup of the model. In sections 2.7.1 and 2.7.2 we focused on two aspects of the model that are potentially responsible for introducing the most numerical error. In section 2.7.1 we showed that the diffusion inherent in the Robert-Asselin filter can be limited by choosing the appropriate value for the filter coefficient. Our 1D experiments yielded a heuristic relationship between the optimal filter coefficient A_{opt}, the gravity wave speed c, and the model timestep dt, namely

$$A_{opt} \propto \frac{c\Delta t}{R_{Earth}}. \tag{5.1}$$

We chose to set $A = 2c\Delta t$ ($R_{Earth} \equiv 1$) and this enables us to choose a filter coefficient that is several orders of magnitude smaller than values typically quoted in the literature.

In section 2.7.2 we examined the error introduced by using 2nd order finite

differences in the latitudinal direction. A series of experiments showed that when the PV field is discontinuous, the order of accuracy of the spectral and 4th order discretisation schemes is reduced to 2nd order. Due to the discontinuous representation of PV used by the CASL model, this fully justifies our use of the more straightforward 2nd order finite difference discretisation in latitude.

In chapter 3 we presented the results of over 150 simulations of freely decaying turbulence on in a shallow fluid layer on a rotating sphere. This is the first time that the CASL algorithm has been applied to such a complex flow in spherical geometry. The novel techniques employed in the algorithm enabled us to explore areas of the $\mathrm{Fr} - \mathrm{Ro}$ parameter space that cannot be reached by other models. This allows us to make a significant contribution to the understanding of flow properties throughout that parameter space.

We used the data generated by our simulations to investigate two important issues in current turbulent flow research: cyclone-anticyclone asymmetry and PV homogenisation. By looking at the skewness of the depth field, we found that away from the geostrophic regime $\mathrm{Fr} \ll 1, \mathrm{Ro} \ll 1$ there is significant cyclone-anticyclone asymmetry, increasing with increasing Ro. This agrees with previous results found by Cho and Polvani [1996]. However, the skewness alone does not give the full picture. By looking at PDFs of the depth field, we demonstrated that a significant 'tail' of strong vortices *of the opposite sign to that predicted by the skewness* can exist. In fact, cyclones, although covering a smaller surface area, tend to be significantly stronger than anticyclones. In many cases they even dominate the flow.

Regarding the second issue, we applied some novel techniques to investigate PV homogenisation and the formation of jets. We found that a statistical approach that measured the clustering of PV contours was hampered by the sensitivity of the clustering schemes, the constraint the PV contours must wrap the sphere, and above all, the necessity of taking longitudinal averages of the PV

contour positions. Instead, we found a local analysis of the palinstrophy field to be more revealing. By correlating the palinstrophy field with the kinetic energy field we were able to compute a local jet velocity and, based on this and the work of Dritschel and McIntyre [2008], a corresponding jet spacing. This provided some confirmation of the theory that there is a critical latitude above which jets will not form. In addition to this, there was an intriguing hint that jet spacing may increase away from the equator but unfortunately our simulations contain too few jets for us to see this clearly.

The first section of chapter 4 focused on the definition of balanced, and hence unbalanced, motion for nonlinear turbulent flows. There are several approaches to this and we chose to examine two. One approach is to set the variables representing the imbalance (or their time derivatives) to zero; the second approach is an iterative procedure that has the novel property that the PV is not approximated and the full equations of motion are solved. We found the iterative OPV balance procedure to be significantly more successful in separating the balanced and unbalanced components of the flow. This result was demonstrated in both physical and spectral space. However, the balanced models worked surprisingly well, even in the equatorial region where the scalings on which they are based are expected to break down. This is important as these balanced models are much faster to compute than the iterative method. Unfortunately, the first order balanced model did not converge for a significant portion of the parameter space (i.e. where the Froude number is large). However, we found that fixing the PV contours in the iterative procedure almost exactly reproduces the results of the first order balanced model (in the region of parameter space where it converges). This enables us to retain the convergence properties of the iterative scheme while significantly reducing the computation time.

Applying the different wave–vortex decompositions to the flows introduced in chapter 3 enabled us to examine the properties of the balanced and unbal-

anced components of these flows. We found that the depth and velocity fields are dominantly balanced with only a *very small* fraction of these fields evolving independently of the PV. Even features in the divergence and acceleration divergence fields can be traced back to features in the PV field. Examining the frequency spectra of the balanced and unbalanced components of the depth field revealed that, when the most accurate decomposition is used, the balanced motion dominates at all scales. The frequency spectra of the balanced and unbalanced components of the divergence field show similar results with the balanced motion dominating across a wide range of scales for three of the four cases we examined. The exceptional case was that which is least representative of geophysical flows. These results show that *there is no clear frequency separation between balanced and unbalanced motion.*

We rounded off the study by considering the impact of the unbalanced motion on the flow. Having found the magnitude of this component (as defined by the OPV balancing routine) to be many orders of magnitude smaller, we investigate how removing it from the simulation affects the properties of the flow. The first discovery was that the impact of the unbalanced motion was, at most, of the same order as the impact of contour surgery. However, if we consider only the times unaffected by contour surgery, we see that the impact of the unbalanced component of the flow is very small and remains so for all time.

An interesting route for further investigation would be to add forcing to the model. This would involve a modification of the way the CASL algorithm deals with PV since the forcing will introduce diabatic effects that the conservative contour advection, by definition, cannot model. Dritschel and Ambaum [1998] and Dritschel and Ambaum [2006] suggest that the best way to accomplish this is to have a grid for the diabatic PV. The diabatic PV is then solved for on the fine grid but transferred to the contours as often as possible in order to take advantage of the conservation properties and fine scale structure inherent

in the contour advection scheme. Dritschel and Fontane have recently developed another method that eliminates the need for a spectral method to solve for the PV. Instead they model the residual PV by a cloud of point vortices whose strengths can be altered to exactly reproduce both the non-conservative forcing and the surgical errors. This new method means that the scheme is no longer restricted by the CFL constraint and there is also no numerical diffusion of the PV field. The inclusion of forcing in the spherical model would provide opportunity to study forced-dissipative flows and complement recent work by Scott and Polvani [2007].

Bibliography

C. Appenzeller, H. C. Davies, and W. Norton. Fragmentation of stratospheric intrusion. *J. Geophys. Res.*, 101:1435–1456, 1996.

M. Arai and T. Yamagata. Asymmetric evolution of eddies in rotating shallow water. *Chaos*, 4:163–175, 1994.

R. Asselin. Frequency filter for time integrations. *Mon. Wea. Rev.*, 100:487–490, 1972.

F. Baer and J. Tribbia. On the complete filtering of gravity modes through nonlinear initialisation. *Mon. Wea. Rev.*, 105:1536–1539, 1997.

J. R. Bates, Y. Li, A. Brandt, S. F. McCormick, and J. Ruge. A global shallow-water numerical model based on the semi-Lagrangian advection of potential vorticity. *Quart. J. Roy. Met. Soc.*, 121:1981–2005, 1995.

J. Cho and L. M. Polvani. The emergence of jets and vortices in freely evolving, shallow-water turbulence on a sphere. *Phys. Fluids*, 8:1531–1552, 1996.

B. Cushman-Roisin and B. Tang. Geostrophic turbulence and emergence of eddies beyond the radius of deformation. *J. Phys. Oceanogr.*, 20:2735–2737, 1990.

S. Danilov and D. Gurarie. Scaling, spectra and zonal jets in beta-plane turbulence. *Phys. Fluids*, 16:2592–2603, 2004.

D. G. Dritschel. Vortex properties of two-dimensional turbulence. *Phys. Fluids*, 5:984–997, 1993.

D. G. Dritschel. On the persistence of non-axisymmetric vortices in inviscid two-dimensional flows. *J. Fluid Mech.*, 371:141–155, 1998.

D. G. Dritschel and M. H. P. Ambaum. A contour-advective semi-Lagrangian numerical algorithm for simulating fine-scale conservative dynamical fields. *Quart. J. Roy. Met. Soc.*, 123:1097–1130, 1997.

D. G. Dritschel and M. H. P. Ambaum. The inclusion of non-conservative forcing into a conservative, contour advection algorithm. *Numerical Methods for Fluid Mechanics VI*, pages 99–110, 1998.

D. G. Dritschel and M. H. P. Ambaum. The diabatic contour advective semi-Lagrangian model. *Mon. Wea. Rev.*, 134:2503–2514, 2006.

D. G. Dritschel and J. Fontane. The hyper-CASL algorithm: a minimally dissipative method for treating subgrid-scale cascades and non-conservative forcing in 2D and layerwise-2D flows. in preparation.

D. G. Dritschel and M. E. McIntyre. Multiple jets as PV staircases: the Phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.*, 2008.

D. G. Dritschel and A. R. Mohebalhojeh. The contour advective semi-Lagrangian algorithm: keeping the balance. *ECMWF Workshop Proceeding on Developments in numerical methods for very high resolution global models.*, 2000.

D. G. Dritschel, L. M. Polvani, and A. R. Mohebalhojeh. The contour-advective semi-Lagrangian algorithm for the shallow water equations. *Mon. Wea. Rev.*, 127:1551–1565, 1999.

D. G. Dritschel and J. Vanneste. Instability of a shallow water potential-vorticity front. *J. Fluid Mech.*, 561:237–254, 2006.

D. G. Dritschel and A. Viúdez. A balanced approach to modelling rotating stably-stratified geophysical flows. *J. Fluid Mech.*, 488:123–150, 2003.

T. Eldevik and K. B. Dysthe. Spiral eddies. *J. Phys. Oceanogr.*, 32:851–869, 2002.

M. Farge and R. Sadourny. Wave-vortex dynamics in rotating shallow water. *J. Fluid Mech.*, 206:433–462, 1989.

R. Fjörtoft. On the changes in the spectral distribution of kinetic energy for two dimensional, non-divergent flow. *Tellus*, 5:225–230, 1953.

R. Ford. The instability of an axisymmetric vortex with monotonic potential vorticity in rotating shallow water. *J. Fluid Mech.*, 280:303–334, 1994a.

R. Ford. The response of a rotating ellipse of uniform potential vorticity to gravity wave radiation. *Phys. Fluids*, 6:3694–3704, 1994b.

B. Fornberg. Numerical study of 2d turbulence. *J. Comp. Phys.*, 25:1–31, 1977.

B. Galperin, H. Nakano, H-P. Huang, and S. Sukoriansky. The ubiquitous zonal jets in the atmospheres of giant planets and Earths oceans. *Geophys. Res. Lett.*, 31, 2004.

A. E. Gill. *Atmosphere-ocean dynamics*. Academic Press, 1982.

G. K. Hakim and A. K. Canavan. Observed cyclone/anticyclone tropopause vortex asymmetries. *J. Atmos. Sci.*, 62:231–240, 2005.

G. K. Hakim, C. Snyder, and D. J. Muraki. A new surface model for cycloneanticyclone asymmetry. *J. Atmos. Sci.*, 59:2405–2420, 2002.

M. Halkidi, Y. Batistakis, and M. Vazirgiannis. On clustering validation techniques. *JIIS*, 17:107–145, 2001.

M. Halkidi and M. Vazirgiannis. Clustering validity assessment: Finding the optimal partitioning of a data set. *Proc. ICDM*, pages 187–194, 2001.

L. J. Heyer, S. Kruglyak, and S. Yooseph. Exploring expression data: Identification and analysis of coexpressed genes. *Genome Research*, 9:1106–1115, 1999.

G. Holloway. Estimation of oceanic eddy transports from satellite altimetry. *Nature*, 323:243–244, 1986.

J. R. Holton. *An introduction to dynamical meteorology.* Academic Press, 1992.

B. J. Hoskins, M. E. McIntyre, and A. W. Robertson. On the use and significance of isentropic potential vorticity maps. *Q. J. Roy. Met. Soc.*, 111:877–940, 1985.

A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31:264–323, 1999.

R. C. Kloosterziel and G. J. F. van Heijst. An experimental study of unstable barotropic vortices in a rotating fluid. *J. Fluid Mech.*, 223:1–24, 1991.

F. Kovàcs, C. Legàny, and A. Babos. Cluster validity measurement techniques. *Proc. of 6th International Symposium of Hungarian Researchers on Computational Intelligence*, 2005.

A. C. Kuo and L. M. Polvani. Nonlinear geostrophic adjustment, cyclone/anticyclone asymmetry, and potential vorticity rearrangement. *Phys. Fluids*, 12:1087–1100, 2000.

C. E. Leith. Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.*, 37:958–968, 1980.

S. K. Lele. Compact finite difference schemes with spectral like resolution. *J. Comp. Phys.*, 103:16–42, 1992.

L. Li, A. P. Ingersoll, A. R. Vasavada, C. C. Porco, A. D. Del Genio, and S. E. Ewald. Life cycles of spots on jupiter from Cassini images. *Icarus*, 172:9–23, 2004.

P. S. Marcus and C. Lee. A model for eastward and westward jets in laboratory experiments and planetary atmospheres. *Phys. Fluids*, 10:1474–1489, 1998.

M. E. McIntyre. The stratospheric poler vortex and sub-vortex: Fluid dynamics and midlatitude ozone loss. *Philos. Trans. Roy. Soc. London*, 352:227–240, 1995.

M. E. McIntyre. Balanced flow. *Encyc. of Atmos. Sci.*, 2:680–685, 2002a.

M. E. McIntyre. Potential vorticity. *Encyc. of Atmos. Sci.*, 2:685–694, 2002b.

M. E. McIntyre and W. A. Norton. Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, 57:1214–1235, 1999.

W. J. McKiver and D. G. Dritschel. Balance in non-hydrostatic rotating stratified turbulence. *J. Fluid Mech.*, 596:201–219, 2008.

J. C. McWilliams. The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.*, 146:21–42, 1984.

J. C. McWilliams. Sub-mesoscale coherent vortices in the ocean. *Rev. Geophys.*, 23:165–182, 1985.

A. R. Mohebalhojeh and D. G. Dritschel. Hierarchies of balance conditions for the f-plane shallow water equations. *J. Atmos. Sci.*, 58:2411–2426, 2000a.

A. R. Mohebalhojeh and D. G. Dritschel. On the representation of gravity waves in numerical models of the shallow water equations. *Q. J. R. Meteorol. Soc*, 126:669–688, 2000b.

A. R. Mohebalhojeh and D. G. Dritschel. Assessing the numerical accuracy of complex spherical shallow water flows. *Mon. Wea. Rev.*, 2007.

M. J. Molemaker, J. C. McWilliams, and I. Yavneh. Baroclinic instability and the loss of balance. *J. Phys. Oceanogr.*, 35:1505–1517, 2005.

W. A. Norton. Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coodinate system and a technique for advecting material contours. *J. Atmos. Sci.*, 51:654–673, 1994.

D. B. Olson. Rings in the ocean. *Ann. Rev. Earth Planet. Sci.*, 19:283–311, 1991.

J. Pedlosky. *Geophysical fluid dynamics*. Springer, 1987.

G. Perret, A. Stegner, M. Farge, and T. Pichon. Cyclone-anticyclone asymmetry of large-scale wakes in the laboratory. *Phys. Fluids*, 18, 2006.

L. M. Polvani, J. C. McWilliams, M. A. Spall, and R. Ford. The coherent structures of shallow water turbulence: Deformation radius effects, cyclone/anticyclone asymmetry and gravity wave formation. *Chaos*, 4:177–186, 1994a.

L. M. Polvani, D. W. Waugh, and R. A. Plumb. On the subtropical edge of the stratospheric surf zone. *J. Atmos. Sci.*, 52:1288–1309, 1994b.

F. J. Poulin and G. R. Flierl. The nonlinear evolution of barotropically unstable jets. *J. Phys. Oceanogr.*, 33:2173–2192, 2003.

W. H. Press, S. A. Teukolski, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in Fortran*. Cambridge University Press, 1992.

P. B. Rhines. Waves and turbulence on a beta-plane. *J. Fluid Mech.*, 69:417–443, 1975.

P. B. Rhines. Jets. *Chaos*, 4:313–339, 1994.

P. B. Rhines and W. R. Young. Homogenization of potential vorticity in planetary gyres. *J. Fluid Mech.*, 122:347–367, 1982.

K. Richards, N. A. Maximenko, F. O. Bryan, and H. Sasaki. Zonal jets in the Pacific ocean. *Geophys. Res. Lett.*, 33, 2006.

H. Ritchie. Application of the semi-lagrangian method to a spectral model of the shallow water equations. *Mon. Wea. Rev.*, 116:1587–1598, 1988.

A. J. Robert. The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Japan*, 44:237–245, 1966.

R. Salmon. *Lectures on geophysical fluid dynamics*. Oxford University Press, 1998.

P. Santangelo, R. Benzi, and B. Legras. The generation of vortices in high resolution, two dimensional decaying turbulence and the influence of initial conditions on the breaking of self similarity. *Phys. Fluids A*, 1:1027–1034, 1989.

R. K. Scott and L. M. Polvani. Forced-dissipative shallow water turbulence on the sphere and the atmospheric circulation of the giant planets. *J. Atmos. Sci.*, 64:3158–3176, 2007.

S. Shetty, X. S. Asay-Davis, and P. S. Marcus. On the interaction of Jupiter's great red spot and zonal jet streams. *J. Atmos. Sci.*, 64:4432–4444, 2007.

A. P. Showman. Numerical simulations of forced shallow-water turbulence: effects of moist convection on the large-scale circulation of jupiter and saturn. *J. Atmos. Sci.*, 2007. in press.

K. S. Smith. Tracer transport along and across coherent jets in two dimensional turbulent flow.. *J.Fluid Mech.*, 544:133–142, 2005.

R. K. Smith and D. G. Dritschel. Revisiting the rossby-haurwitz wave test case with contour advection. *J. Comp. Phys.*, 217:473–484, 2006.

C. Snyder, D. J. Muraki, R. Plougonven, and F. Zhang. Inertia-gravity waves generated within a dipole vortex. *J. Atmos. Sci.*, 2007.

J. Sommeria, S. D. Meyers, and H. L. Swinney. Laboratory model of a planetary eastward jet. *Nature*, 337:58–61, 1989.

A. Stegner and D. G. Dritschel. A numerical investigation of the stability of isolated vortices beyond the quasi-geostrophic regime. *J. Phys. Oceanogr.*, 30: 2562–2573, 2000.

C. Temperton and A. Staniforth. An efficient two-time level semi-Lagrangian semi-implicit integration scheme. *Quart. J. Roy. Met. Soc.*, 123:1888–1903, 1987.

J. Theiss. Equatorward energy cascade, critical latitude and the predominance of cyclonic vortices in geostrophic turbulence. *J. Phys. Oceanogr.*, 34:1663 – 1678, 2004.

J. Thuburn. A PV based shallow water model on a hexagonal-icosahedral grid. *Mon. Wea. Rev.*, 125:2328–2347, 1997.

A. R. Vasavada and A. P. Showman. Jovian atmospheric dynamics: an update after Galileo and Cassini. *Rep. Prog. Phys.*, 68:1935–1996, 2005.

J. Venn. The law of error. *Nature*, 36:411–412, 1887.

A. Viúdez and D. G. Dritschel. Dynamic potential vorticity initialization and the diagnosis of mesoscale motion. *J. Phys. Oceanogr.*, 34:2761–2773, 2004a.

A. Viúdez and D. G. Dritschel. Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.*, 521:343–352, 2004b.

A. Viúdez and D. G. Dritschel. Spontaneous generation of inertia-gravity wave packets by balanced geophysical flows. *J. Fluid Mech.*, 553:107–117, 2006.

D. W. Waugh and R. A. Plumb. Contour advection with surgery: A technique for investigating finescale structure in tracer transport. *J. Atmos. Sci.*, 51: 530–540, 1994.

M. C. Wheeler. Equatorial waves. *Tropical Meteorology*, 2002.

A. A. White. A view of the equations of meteorological dynamics and various approximations. In J. Norbury and I. Roulstone, editors, *Large-scale atmosphere-ocean dynamics I: Analytical methods and numerical models.*, pages 1–100. Cambridge University Press, 2002.

P. D. Williams, P. L. Read, and T. W. N. Haine. Spontaneous generation and impact of inertia-gravity waves in a stratified, two-layer flow. *Geophys. Res. Lett.*, 30:2255–2257, 2003.

V. Wirth. Cyclone-anticyclone asymmetry concerning the height of the thermal and the dynamical tropopause. *J. Atmos. Sci.*, 58:26–37, 2000.

A. Youssef and P. S. Marcus. The dynamics of jovian white ovals from formation to merger. *Icarus*, 162:74–93, 2003.